

A formal framework to define margins in industrial processes

Adrien Touboul, Julien Reygner, Fabien Mangeant, Pierre Benjamin

▶ To cite this version:

Adrien Touboul, Julien Reygner, Fabien Mangeant, Pierre Benjamin. A formal framework to define margins in industrial processes. 2019. hal-02156493

HAL Id: hal-02156493 https://hal.science/hal-02156493

Preprint submitted on 14 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A formal framework to define margins in industrial processes.

Adrien Touboul^{1,2}, Julien Reygner¹, Fabien Mangeant^{2,3}, and Pierre Benjamin^{2,4}

¹Université Paris-Est, CERMICS (ENPC), 77455 Marne-la-Vallée, France ²IRT SystemX, Paris-Saclay, France ³RENAULT ⁴AIRBUS

Abstract

A common engineering pattern to handle uncertainties in industrial processes is to use margins, understood as an amount of something included so as to be sure of success or safety. It is supported by practices which differ depending on the discipline. This paper proposes a formal framework to describe these practices independently from the discipline considered, by introducing the concept of model of margin. The structure of the model of margin identifies the common steps between these practices. This framework can be used to formulate some general engineering questions on margins as mathematical problems, and two examples of such questions are developed. Some classical notions of margin from the literature are recasted in this framework as their associated models of margin are being constructed.

1 Introduction

1.1 Motivation: decision in industrial processes

Uncertainties are inherently generated during the design, manufacturing and operation of large complex systems. For example, very classical ones such as uncertainties in manufacturing or uncertainties in metrology are well documented and managed by appropriate standards and referentials [8, 18]. Other sources of uncertainty are well identified but less standardized, except some in specific industrial fields such as the lack of knowledge related to the maturity of a design [3].

In order to sketch a picture of the origin of these uncertainties, one has to take into account the complexity of an industrial organization at different levels:

- Multiple stakeholders with diverse interests are interacting together in the ecosystem;
- Multiple disciplines assess and predict the performance of the system at different levels;
- Customers' expectations and business requests are variable during the life-cycle of a product;
- The quality and availability of the engineering resources can limit the operations.

As a matter of fact, it is impossible to get a fully detailed real-time picture of the performances. That is why industrial organizations have been shaped to manage this risk and each stakeholder takes its own risk. At a first level, we can define uncertainties as things that are not known, or known only imprecisely. We assume that the uncertainties are factual and identifiable: things are either known, unknown or known to a quantifiable degree. In this first approach, we can differentiate [23]:

- The unknown unknowns, that are the things that were not anticipated;
- The known unknowns, that are the things that were already known from past experience but which ranges were not certain at the time of the analysis.

The deployment of Uncertainty Quantification (UQ) methods has been done in pioneering fields, in order to manage the known unknowns (or uncertainties) in a more quantitative way. However, it is often perceived as an additional cost in a disciplinary scope [22]. In order to overcome this difficulty, we propose to focus on a more operating concept in engineering, the margin. Margin is the cornerstone of the system uncertainty management, as it encompasses the knowledge of the system boundaries, the risk management at an industrial program level and the mathematical modeling of the concept.

The ultimate motivation behind this work is that a margin should be more easily used by decision makers to handle the known unknowns or the unknown unknowns. For example, an usual way to make a decision in an uncertain context is to "take a margin", but the margin is often either implicitly defined or focusing on a very specific aspect of the system (such as the gain margin in control). The work in [6] identifies a consequence of this lack of formalism, namely the difficulty to avoid the overdesign in complex system design. One of the main purposes of this paper is to propose an abstract formalization of the concept of margin, independent from the domain or discipline. This framework can then be used to formulate some general engineering questions on margins as mathematical problems.

We base our definition of a margin on an extensive description with well-defined mathematical objects. Thus, the model of margin developed in this paper might not be practical *per se* for decision making. However, it can help the engineer interested in developing margin practices at a large scale and margin computations in nonlinear models, for instance by rigorously formulating the questions: *How much margin is there? How can I take a margin m?* or *What is the impact on my system when I take a margin m?* It is then a complementary approach with other formalizing works such as those in [7, 12, 13, 30].

1.2 The concept to be formalized

The word "margin" has numerous meanings in engineering, it is thus necessary to precise what is being formalized hereafter. The concept can generally be described as an amount of something included so as to be sure of success or safety (Oxford Dictionary). It includes for instance the concept of performance margin defined in the NASA technical report [4, pp 98], a performance margin is defined as the distance from the achieved value of a performance measure at any point in time to a decision boundary for that measure. In a system design context, its impact is described in [23] as designing systems to be more capable, to withstand worse environments, and to last longer than "necessary". In a safety context [32] it can be formulated as something that is over and above what is strictly necessary and that is designed to provide for emergencies [...]. In other contexts, it can be called safety margin, margin to damage, flexibility margin, and so on.

The common factor of each of these concepts is that some uncertainties are taken into account - they might be implicit or explicit, known unknowns and even unknown unknowns. A loss (or opportunity) in a certain sense is associated to these uncertainties. Finally, a quantity is taken or *reserved* to prevent the loss (or ensure the opportunity) associated to these uncertainties. Another commonly shared factor is that the margin is a one-dimensional deterministic value that relates to a meaningful quantity. We chose to keep this aspect as it seems important to support decision-making. In all the examples that we encountered in our literature review (some are presented in Section 4), we have been able to formalize each concept with a model of margin as long as it has the above-mentioned characteristics.

What is not being formalized The concept of profit margin, which is the gain made by a given operation, is not formalized in this paper. The margin of error, which has a specific meaning in Statistics as the range of the confidence interval, is also out of the scope of our work. Last, this work does not aim specifically at studying the margin in a machine learning context - the margin being understood as the distance to a decision boundary in an algorithm.

1.3 Margin: both actual value and requirement

In some cases, one is only interested in defining and choosing the value of the margin requirement without thinking of the choice of a particular state of the system/phenomenon described on which there would be margins. For instance, one could need to define a design space with margin without choosing any design point *a priori*. Even if it is not always explicit, we think that these practices can be decomposed into two steps:

- 1. Defining how to compute the margin for a given state
- 2. Expressing a requirement on the margin.

The requirement is then that the state describing the phenomenon must have a greater margin than the required margin, computed with the provided definition. Properly defining how to compute a margin for a given state (Step 1) is the topic of Section 2. Choosing the required margin (Step 2) is part of what we call the margin calibration and is mentioned in Section 5.1.

In the literature, the word margin can be used to refer to the actual value for a given state or to the requirement. We then call *effective margin* the actual value, the extra quantity that can be calculated or estimated by computer or real-life experiments. The *demanded margin* is the requirement made on this extra quantity.

1.4 Outline of the paper

In Section 1, we introduce the industrial stakes that motivated this work and describe informally the concept of margin, which is the topic of this paper. The formal definitions of the model of margin and of the margin are given in Section 2. A precision on the definition of the acceptance sets with risk measures is provided in Section 3. Our work is linked with the literature in Section 4, by recasting some existing practices with models of margins. Section 5 outlines how some classes of engineering problems, namely the margin calibration and margin allocation, can be formulated thanks to models of margins. Section 6 concludes this work.

1.5 Notation

- The complement of a set \mathscr{A} is denoted by \mathscr{A}^c .
- If $x \in \mathbb{R}$ then $[x]^+ = \max(x, 0)$.
- The quantile of order r of a real-valued random variable X is denoted by $q_r(X)$. If X is a standard normal variable, its quantile is simply denoted by φ_r .
- We denote the set of integers between 1 and n by [1; n].

2 Definitions

The aim of this section is to provide a rigorous framework to describe, define, interpret and compute a margin independently from any field considered. We stated in the introduction that the investigated margin concept is *an extra quantity being taken*. Our thesis is that this extra quantity can be computed as a distance to a reference. However, this reference is not always a critical point to avoid or a point of performance to aim at and the distance does not always include all the variables of the problem.

2.1 Model of margin

A model of margin represents the information required from the analysis of the phenomenon to uniquely define a computed margin. Hereafter are presented two cases, deterministic and random, to put in light the fact that a margin can be defined without any probabilistic consideration, although the random case can be seen as a generalization of the deterministic one.

2.1.1 Deterministic case

Let us consider a variable of interest, denoted by u, for which one is interested in computing a margin. If u is deterministic and takes its values in a state space E, the model of margin is composed of:

Definition 2.1 (Problem description - deterministic case). The *problem description* is defined by the triple $(E, \mathcal{C}, \mathcal{A})$ with \mathcal{C} and \mathcal{A} two subsets of the set E.

 \mathscr{C} is the set of the **problem constraints**. It characterizes the values that can be taken by the variables describing the phenomenon u due to constraints given by the model - such as physical or logical constraints. \mathscr{A} is the **acceptance set**. It discriminates the values of u which have an acceptable risk and those that do not have.

Definition 2.2 (Probing set - deterministic case). Let $(E, \mathcal{C}, \mathcal{A})$ be a problem description. The family of *probing sets* is a family of subsets of E indexed by $u \in \mathcal{C}$, and is denoted by $(\mathcal{G}_u)_{u \in \mathcal{C}}$. The intersection between a probing set and the problem constraints is $\mathcal{G}_{u|\mathcal{C}} = \mathcal{G}_u \cap \mathcal{C}$.

The **probing set** \mathcal{G}_u contains the possible scenarios that one wants to consider, namely, the values of the variable of interest along which one wants to compute a margin.

Definition 2.3 (Coordinates of interest - deterministic case). Let $(E, \mathcal{C}, \mathcal{A})$ be a problem description and $(\mathcal{G}_u)_{u \in \mathscr{C}}$ a family of probing sets. A set of *coordinates of interest* is a metric space $(\mathbf{S}, d_{\mathbf{S}})$, provided with a family of functions $\phi_u : \mathcal{G}_u \to \mathbf{S}$ indexed by $u \in \mathscr{C}$. The family $(\phi_u)_{u \in \mathscr{C}}$ is called the family of *coordinate functions*.

Given $s \in \mathbf{S}$ and $B \subset \mathbf{S}$, we denote $d_{\mathbf{S}}(s, B) = \inf_{t \in B} d_{\mathbf{S}}(s, t)$. If $B = \emptyset$, $d_{\mathbf{S}}(s, B) = +\infty$ by convention. For modeling purposes, it may be useful to allow $d_{\mathbf{S}}$ to take the value $+\infty$, when $\mathbf{S} = \mathbb{R} \cup \{+\infty, -\infty\}$ for instance.

The set of coordinates of interest **S** represents the quantities on which one wants to express a margin, which will be given in terms of the distance $d_{\mathbf{S}}$. These coordinates can be a subset of the coordinates used to describe the elements of the state space E but could also possibly be some coordinates of the probing sets of the family $(\mathcal{G}_u)_{u \in \mathscr{C}}$. ϕ_u is the coordinate function that associates each element of \mathcal{G}_u with its coordinates in **S**.

Definition 2.4 (Model of margin - deterministic case). A model of margin is the combination of a problem description, a family of probing sets and a set of coordinates of interest with the corresponding coordinate functions. It is denoted by:

$$\mathsf{M} = (\mathsf{E}, \mathscr{C}, \mathscr{A}, (\mathscr{G}_u)_{u \in \mathscr{C}}, \mathbf{S}, d_{\mathbf{S}}, (\phi_u)_{u \in \mathscr{C}})$$
(1)

A complete example will be given in Section 2.1.3. Before, the notion of margin is defined in Section 2.1.2.

2.1.2 Margin

As discussed in the introduction, we look for an expression of the margin as a measure of how far the state of the system, u, is from some risky states. The risky states are given by the complement \mathscr{A}^c of the acceptance set and the range of the states investigated is given by the probing set $\mathscr{G}_{u|\mathscr{C}}$. Thus, the risky states - i.e the set of the probed non acceptable variables of interest - are in $\mathscr{G}_{u|\mathscr{C}} \cap \mathscr{A}^c$. This motivates our definition:

Definition 2.5 (Margin). Let $\mathsf{M} = (\mathsf{E}, \mathscr{C}, \mathscr{A}, (\mathscr{G}_u)_{u \in \mathscr{C}}, \mathbf{S}, d_{\mathbf{S}}, (\phi_u)_{u \in \mathscr{C}})$ be a given margin model and $u \in \mathscr{C}$ a variable of interest. The *margin* at the point u for the model M is defined as:

$$\mathbf{m}\left(u;\mathbf{M}\right) = \begin{cases} d_{\mathbf{S}}\left(\phi_{u}(u),\phi_{u}(\mathscr{A}^{c}\cap\mathscr{G}_{u|\mathscr{C}})\right) & \text{if } u\in\mathscr{A}, \\ -d_{\mathbf{S}}\left(\phi_{u}(u),\phi_{u}(\mathscr{A}^{c}\cap\mathscr{G}_{u|\mathscr{C}})^{c}\cap\phi_{u}(\mathscr{G}_{u|\mathscr{C}})\right) & \text{if } u\notin\mathscr{A}. \end{cases}$$
(2)

The computed margin $\mathbf{m}(u; \mathbf{M})$ (Section 1.3) can be interpreted as the margin at the state u on the coordinates represented by \mathbf{S} aggregated with $d_{\mathbf{S}}$, for the acceptance set \mathcal{A} , the phenomenon being described by the problem constraints and equations in \mathcal{C} , and considering only the evolutions of the state in \mathcal{G}_u . The model of margin describes how to compute margins, it is then relevant to compare the margins of two different states only if they have been computed with the same model of margin.

Intuitively, the margin for a given value u of a variable of interest is the distance from this point to the acceptance set, for some chosen scenarios and focusing only on some coordinates describing the variable of interest. The core property of the margin, implied by (2) is:

Proposition 2.1 (Acceptability property of the margin). If $u' \in \mathcal{G}_{u|\mathscr{C}}$ and $d_{\mathbf{S}}(\phi_u(u), \phi_u(u')) < \mathbf{m}(u; \mathsf{M})$ then $u' \in \mathscr{A} \cap \mathscr{C}$.

If the margin m is positive, this means that the coordinates of u in \mathbf{S} , which values are $\phi_u(u)$, are at least at a distance m from any possible unacceptable value explored in $\mathscr{G}_{u|\mathscr{C}}$. Another way to express it is that if the system evolves in $\mathscr{G}_{u|\mathscr{C}}$, the coordinates would have to travel a distance of at least m to reach an unacceptable value. In that sense, a margin can be interpreted as an "extra quantity taken" to ensure success or safety.

Now, let us assume that the probing sets $\mathscr{G}_{u|\mathscr{C}}$ and the coordinates function ϕ_u are the same for every point $u \in \mathscr{C}$, and respectively equal to $\mathscr{G}_u = \mathscr{C}$ and a given function ϕ .

Proposition 2.2 (Bounds on the margin increase). Let $(u, v) \in \mathscr{C}^2$.

- (1) If u and v both belong to \mathscr{A} or both belong to \mathscr{A}^c , we have: $|\mathbf{m}(u; \mathbf{M}) \mathbf{m}(v; \mathbf{M})| \leq d_{\mathbf{S}}(\phi(v), \phi(u))$.
- (2) If $u \in \mathscr{A}^c$ and $v \in \mathscr{A}$, we have: $\mathsf{m}(v;\mathsf{M}) \leq [\mathsf{m}(u;\mathsf{M}) + 2d_{\mathbf{S}}(\phi(u),\phi(v))]^+$.

Let us assume in addition that **S** is a normed vector space, $d_{\mathbf{S}}$ is induced by the norm and $\phi(\mathscr{C})$ is convex. We can refine the result to:

- (3) If $u \in \mathscr{A}^c$ and $v \in \mathscr{A}$, we have: $\mathsf{m}(v;\mathsf{M}) \leq [\mathsf{m}(u;\mathsf{M}) + d_{\mathbf{S}}(\phi(u),\phi(v))]^+$.
- (4) If furthermore $\phi^{-1}(\phi(\mathscr{A})) = \mathscr{A}$, then for any $(u, v) \in \mathscr{C}^2$: $|\mathsf{m}(v; \mathsf{M}) \mathsf{m}(u; \mathsf{M})| \le d_{\mathbf{S}}(\phi(u), \phi(v))$.

Proof: Assertion (1) comes directly from the Lipschitz property of a distance: $\forall B \subset \mathbf{S}, |d_{\mathbf{S}}(x, B) - d_{\mathbf{S}}(y, B)| \leq d_{\mathbf{S}}(x, y).$

Let us consider the case $u \in \mathscr{A}^c$ and $v \in \mathscr{A}$ of Assertion (2). If $\phi(v) \in \phi(\mathscr{A}^c \cap \mathscr{C})$ then $\mathsf{m}(v;\mathsf{M}) = 0$. If $\phi(v) \notin \phi(\mathscr{A}^c \cap \mathscr{C})$, then $\phi(v) \in \phi(\mathscr{A}^c \cap \mathscr{C})^c \cap \phi(\mathscr{C})$, we have:

$$-\mathbf{m}\left(u;\mathbf{M}\right)+\mathbf{m}\left(v;\mathbf{M}\right)=d_{\mathbf{S}}\left(\phi(u),\phi(\mathscr{A}^{c}\cap\mathscr{C})^{c}\cap\phi(\mathscr{C})\right)+d_{\mathbf{S}}\left(\phi(v),\phi(\mathscr{A}^{c}\cap\mathscr{C})\right).$$

Since $\phi(v) \in \phi(\mathscr{A}^c \cap \mathscr{C})^c \cap \phi(\mathscr{C})$ and $\phi(u) \in \phi(\mathscr{A}^c \cap \mathscr{C})$, we can bound both distances by $d_{\mathbf{S}}(\phi(u), \phi(v))$. We have then $-\mathbf{m}(u; \mathsf{M}) + \mathbf{m}(v; \mathsf{M}) \leq 2d_{\mathbf{S}}(\phi(u), \phi(v))$, which proves Assertion (2). The equality occurs when v is the arginf of $d_{\mathbf{S}}(\phi(u), \phi(\cdot))$ and u is the arginf of $d_{\mathbf{S}}(\phi(\cdot), \phi(v))$.

If $\phi(\mathscr{C})$ is convex, $d_{\mathbf{S}}$ is induced by a norm $\|\cdot\|_{\mathbf{S}}$ and $\phi(v) \in \phi(\mathscr{A}^c \cap \mathscr{C})^c \cap \phi(\mathscr{C})$, it is possible to be more precise.

Let $t_v = \inf \{t \in [0,1], (1-t)\phi(v) + t\phi(u) \in \phi(\mathscr{A}^c \cap \mathscr{C})\}$ and $t_u = \inf \{t \in [0,1], (1-t)\phi(u) + t\phi(v) \in \phi(\mathscr{A}^c \cap \mathscr{C})^c\}$. By convexity of $\phi(\mathscr{C})$ $(1-t)\phi(u) + t\phi(v) \in \phi(\mathscr{C})$, for all $t \in [0,1]$. The continuity of the distance gives us:

$$\mathbf{m}(v;\mathbf{M}) = d_{\mathbf{S}}(\phi(v),\phi(\mathscr{A}^c \cap \mathscr{C})) \le d_{\mathbf{S}}(\phi(v),(1-t_v)\phi(v) + t_v\phi(u)) = t_v \|\phi(u) - \phi(v)\|_{\mathbf{S}},\tag{3}$$

and

$$-\mathbf{m}(u;\mathbf{M}) = d_{\mathbf{S}}(\phi(u),\phi(\mathscr{A}^c \cap \mathscr{C})^c \cap \mathscr{C}) \le d_{\mathbf{S}}(\phi(u),(1-t_u)\phi(u) + t_u\phi(v)) = t_u \|\phi(u) - \phi(v)\|_{\mathbf{S}}$$
(4)

As $t_u + t_v \leq 1$, we deduce the following inequality:

$$0 \le \mathsf{m}(v;\mathsf{M}) - \mathsf{m}(u;\mathsf{M}) \le (t_u + t_v) \|\phi(u) - \phi(v)\|_{\mathbf{S}} \le d_{\mathbf{S}}(\phi(u),\phi(v))$$

This proves Assertion (3).

Last, if $\phi^{-1}(\phi(\mathscr{A})) = \mathscr{A}$, if $v \in \mathscr{A}$ we always have $\phi(v) \in \phi(\mathscr{A}^c \cap \mathscr{C})^c \cap \phi(\mathscr{C})$. By the proof of Assertion (3), $0 \leq \mathsf{m}(v;\mathsf{M}) - \mathsf{m}(u;\mathsf{M}) \leq d_{\mathbf{S}}(\phi(v),\phi(u))$ always holds in that case and thus $|\mathsf{m}(v;\mathsf{M}) - \mathsf{m}(u;\mathsf{M})| \leq d_{\mathbf{S}}(\phi(u),\phi(v))$. The latter formula being symmetric in u and v, it also holds when $u \in \mathscr{A}$ and $v \in \mathscr{A}^c$. Using Assertion (1) we have proved this inequality for all $(u,v) \in \mathscr{C}^2$.

Proposition 2.2 gives an interpretation of a negative margin: if a point u has a negative margin m, then any point v with a positive margin has coordinates in **S** at least at a distance |m| from $\phi(u)$, under the assumption of Assertion (4) in Proposition 2.2. Thus, instead of quantifying how far a point is from coordinates of acceptable points, the negative margin monitors how far the point is from having a positive margin and is more robust in a sense. That is why $\phi_u(\mathscr{A}^c \cap \mathscr{G}_{u|\mathscr{C}})^c \cap \phi_u(\mathscr{G}_{u|\mathscr{C}})$ has been chosen instead of the "simple choice" $\phi_u(\mathscr{A} \cap \mathscr{G}_{u|\mathscr{C}})$ in Definition 2.5.

2.1.3 Example of margins: setup of a water heater

Let us illustrate our example of definitions of margins through a simple problem. In the setup of a water heater, the user can choose two parameters: the maximum amount of water V and the power of the heating resistance P. Another variable of interest is the heating time T. We can then construct the state space $E = \{u = (P, V, T) \in \mathbb{R}^3\}$.

The specifications of the water heater give limits on these quantities.

- The maximum amount of water cannot exceed the volume of the storage $V_{\text{max}} > 0$ and it is nonnegative.
- The power must also be nonnegative and the system can only operate up to a given power $P_{\text{max}} > 0$.
- The heating time is linked to the power and volume by the model $T = G_P^V$ with G being the product of the heating capacity of the water and the increase of temperature required. It is assumed to be a known constant.

The set of the problem constraints is then :

$$\mathscr{C} = \left\{ (P, V, T) \in \mathbf{E} \left| 0 \le P \le P_{\max} 0 \le V \le V_{\max}, T = G \frac{V}{P} \right\}.$$

The user also provides his or her own constraints for a state to be acceptable. The heating time T of the tank at full load must be lower than $T_{\text{lim}} > 0$ and the volume must be greater than a conservative estimation of their needs $V_{\text{needs}} > 0$. The acceptance set is constructed from these constraints:

$$\mathscr{A} = \{ (P, V, T) \in \mathcal{E} \mid V > V_{\text{needs}}, T < T_{\text{lim}} \}.$$

An illustration of the problem description $(E, \mathcal{C}, \mathcal{A})$ is provided in Figure 1.



Figure 1: Illustration of the problem description $(E, \mathcal{C}, \mathcal{A})$ in the plane (P, V) with the numerical values $V_{\text{needs}} = 37 \text{ L}, V_{\text{max}} = 200 \text{ L}, P_{\text{max}} = 30 \text{ kW}, T_{\text{lim}} = 30 \text{ min and } G = 1.05 \text{ min kW L}^{-1}.$

With this problem description, two types of scenarios are investigated.

Scenario 1 The first set of scenarios corresponds to a decrease in power while keeping a fixed volume. The associated family of probing sets for $u = (P, V, T) \in \mathcal{C}$ is:

$$\mathscr{G}_u = \{ (P', V', T') \in \mathbf{E} \mid V' = V, P' < P \}$$

For this family of probing sets, if $u = (P, V, T) \in \mathscr{A}$ then a triple $(P', V', T') \in E$ belongs to $\mathscr{A}^c \cap \mathscr{G}_{u|\mathscr{C}}$ if and only if V' = V, $0 \leq P' \leq \frac{GV}{T_{\lim}}$ and $T' = \frac{GV}{P'}$. With these probing sets, a first possibility is to compute a margin on P, i.e choosing P as the coordinate

With these probing sets, a first possibility is to compute a margin on P, i.e choosing P as the coordinate of interest. In that case, the set of coordinates of interest is $\mathbf{S} = \mathbb{R}$ with the distance $d_{\mathbf{S}}(x,y) = |x - y|$. The coordinate functions are the projections defined by $\phi_u(P', V', T') = P'$, for all u in \mathscr{C} . The projection of the non acceptable coordinates in the probing set is then:

$$\phi_u\left(\mathscr{A}^c\cap\mathscr{G}_{u|\mathscr{C}}\right) = \left[0; G\frac{V}{T_{\lim}}\right).$$

If we denote by M_P the margin model in this context, for $u = (P, V, T) \in \mathcal{A}$, the margin on the power is:

$$\mathbf{m}(u;\mathbf{M}_P) = d_{\mathbf{S}}(\phi_u(u),\phi_u\left(\mathscr{A}^c \cap \mathscr{G}_{u|\mathscr{C}}\right)) = d_{\mathbf{S}}\left(P,\left[0;G\frac{V}{T_{\lim}}\right)\right) = P - G\frac{V}{T_{\lim}}.$$
(5)

As a consequence, if the user wants to decrease the power consumed by the water heater, he or she knows that the maximum acceptable decrease is at least $\mathsf{m}(u;\mathsf{M}_P) = P - G \frac{V}{T_{\lim}}$, as long as the evolution stays in $\mathscr{G}_{u|\mathscr{C}}$ - i.e the volume of water does not change. This margin is illustrated in Figure 2.

Let us look now at the margin on the heating time T. We consider as a family of coordinate functions the projection on the heating time axis $\phi_u(P', V', T') = T'$. The projection of the non acceptable coordinates is now $\phi_u\left(\mathscr{A}^c \cap \mathscr{G}_{u|\mathscr{C}}\right) = (T_{\lim}; +\infty)$. Let us denote by M_T the margin model for a margin on T. The margin at the point $u = (P, V, T) \in \mathscr{A}$ is :

$$\mathbf{m}\left(u;\mathbf{M}_{T}\right) = d_{\mathbf{S}}\left(\phi_{u}(u),\phi_{u}\left(\mathscr{A}^{c}\cap\mathscr{G}_{u|\mathscr{C}}\right)\right) = d_{\mathbf{S}}\left(T,\left(T_{\mathrm{lim}};+\infty\right)\right) = T_{\mathrm{lim}} - T.$$
(6)

This margin tells us that any increase in the heating time under $T_{\text{lim}} - T$ would still satisfy the constraints, as long as the point considered is in $\mathscr{G}_{u|\mathscr{C}}$. In fact, one can see that any increase in the heating time over $T_{\text{lim}} - T$ would not be acceptable by definition, but this information is not provided by the margin.

Finally, our framework also allows to compute a margin on the volume with the coordinate function $\phi_u(P', V', T') = V'$. However, in that case, the projection of the non acceptable coordinates is: $\phi_u\left(\mathscr{A}^c \cap \mathscr{G}_{u|\mathscr{C}}\right) = \{V\}$. The margin is then $d_{\mathbf{S}}(\phi_u(u), \{V\}) = d_{\mathbf{S}}(V, \{V\}) = 0$. The meaning of a zero margin is consistent with Proposition 2.1. It is impossible to exhibit an open interval centered on V such that each point of $\mathscr{G}_{(P,V,T)|\mathscr{C}}$ that is not acceptable - i.e that is also in \mathscr{A}^c - has a volume outside this interval.



Figure 2: Margin on the power in Scenario 1. P = 11 kW and V = 112 L.

Scenario 2 Let us now consider that the user has the possibility to change P and V in order to minimize their cost:

$$C(P,V) = c_0 PV,$$

where P is expressed in kW, V is expressed in L and c_0 in $\in kW^{-1}L^{-1}$. Starting from a point $u = (P, V, T) \in \mathscr{C} \cap \mathscr{A}$, the probing set $\mathscr{G}_{u|\mathscr{C}}$ is constructed as the continuous limit of a sequence of successive minimizers of

the functional C(P, V) with small steps in the space of pairs (P, V). To define it precisely, we first provide this space with a norm $\|\cdot\|_{\beta_P,\beta_V}$ defined by the identity:

$$\|(P,V)\|_{\beta_P,\beta_V}^2 = \left(\frac{P}{\beta_P}\right)^2 + \left(\frac{V}{\beta_V}\right)^2,$$

where β_P and β_V are positive constants, respectively expressed in kW and L. Let us now define the sequence of points $u_n = (P_n, V_n, T_n) \in E$ by $u_0 = u$ and for all $n \ge 0$,

$$(P_{n+1}, V_{n+1}) = \operatorname*{argmin}_{\|(P-P_n, V-V_n)\|_{\alpha,\beta}^2 \le \epsilon \|(P_n, V_n)\|_{\beta_P, \beta_V}^2} C(P, V)$$

and $T_{n+1} = G_{\frac{P_{n+1}}{P_{n+1}}}^{V_{n+1}}$. In the $\epsilon \to 0$ limit, this sequence is approximated by the continuous curve $t \to u(t) = (P(t), V(t), T(t)) \in \mathcal{C}$ defined by:

$$\dot{P}(t) = -\frac{\beta_P}{\beta_V}V(t), \qquad \dot{V}(t) = -\frac{\beta_V}{\beta_P}P(t), \qquad T(t) = G\frac{V(t)}{P(t)}.$$
(7)

The solution of the coupled system of ODEs for (P(t), V(t)) is $P(t) = P \cosh(-t) + V \frac{\beta_P}{\beta_V} \sinh(-t)$ and $V(t) = V \cosh(-t) + P \frac{\beta_V}{\beta_P} \sinh(-t)$, which finally leads to the probing set $\mathcal{G}_u = \{u(t) = (P(t), V(t), T(t)), t \ge 0\}$. One can verify that it can be rewritten:

$$\mathscr{G}_{u|\mathscr{C}} = \left\{ \left(P', V', G\frac{V'}{P'}\right) \in \mathscr{C} \left| \left(\frac{P'}{\beta_P}\right)^2 - \left(\frac{V'}{\beta_V}\right)^2 = \left(\frac{P}{\beta_P}\right)^2 - \left(\frac{V}{\beta_V}\right)^2, P' \le P \right\}.$$

In this case, illustrated by Figure 3, the margin on the power is the difference between the power Pand the power of the point where the curve crosses the line $V = V_{\text{needs}}$, namely $\sqrt{P^2 - \frac{\beta_P^2}{\beta_V^2}(V^2 - V_{\text{needs}}^2)}$. The margin on the power is thus $\mathbf{m}(u; \mathbf{M}_P) = P - \sqrt{P^2 - \frac{\beta_P^2}{\beta_V^2}(V^2 - V_{\text{needs}}^2)}$. The margin on the volume is

simply $\mathbf{m}(u; \mathbf{M}_V) = V - V_{\text{needs}}.$

One can verify that the heating time is decreasing all along the optimization curve. The margin on the heating time is then the difference between T and the heating time of the point where the curve crosses $V = V_{\text{needs}}$. Its value is $\mathsf{m}(u;\mathsf{M}_T) = T - G \frac{V_{\text{needs}}}{\sqrt{P^2 - \frac{\beta_P^2}{\beta_V^2}} (V^2 - V_{\text{needs}}^2)}$. The interpretation of such a margin is

that any decrease in the heating time along the optimization curve $\mathscr{G}_{u|\mathscr{C}}$ smaller than $\mathsf{m}(u;\mathsf{M}_T)$ leads to an acceptable state.

2.1.4 Random case

When the phenomenon is modeled by a random variable U taking its values in a measurable space E, defined on a probability triple $(\Omega, \mathcal{T}, \mathbf{P})$, the definitions of Section 2.1.1 need to be slightly modified. The set of E-valued random variables is denoted by \mathbb{L}^0 (Ω, \mathbb{E}) .

Definition 2.6 (Problem description - random case). The *problem description* is defined by the triple $(E, \mathcal{C}, \mathcal{A})$ with \mathcal{C} and \mathcal{A} two subsets of $\mathbb{L}^0(\Omega, E)$.

Now, \mathscr{C} and \mathscr{A} contain random variables. In practice, this means that \mathscr{C} can also contain conditions on the possible probability distributions of U, provided by the modeling choices. The acceptance set \mathscr{A} may also depend on the distribution of U.

Definition 2.7 (Probing set - random case). Let $(E, \mathcal{C}, \mathcal{A})$ be a problem description. The *family of probing* sets is a family of subsets of $\mathbb{L}^0(\Omega, E)$ indexed by U in \mathcal{C} and is denoted by $(\mathcal{G}_U)_{U \in \mathcal{C}}$. The set of random variables in the probing set satisfying the problem constraints is denoted by $\mathcal{G}_{U|\mathcal{C}} = \mathcal{G}_U \cap \mathcal{C}$.

One of the main interests in defining a model of margin with random variables lies in the possibility of describing the probing set as a subset of the random variables of interest on E and not only on some of their statistical indicators. One could define for instance the set of the random variables almost surely greater than a given random variable as a probing set.



Figure 3: Margin on the power and volume in Scenario 2. The probing set follows an optimization curve starting at P = 19 kW, V = 150 L and $\frac{\beta_V}{\beta_P} = 10 \text{ L kW}^{-1}$.

Definition 2.8 (Coordinates of interest - random case). Let $(E, \mathcal{C}, \mathcal{A})$ be a problem description and $(\mathcal{G}_U)_{U \in \mathscr{C}}$ be a family of probing sets. A set of coordinates of interest is a metric space $(\mathbf{S}, d_{\mathbf{S}})$, provided with a family of functions $\phi_U : \mathcal{G}_U \to \mathbf{S}$ with $U \in \mathscr{C}$. The family $(\phi_U)_{U \in \mathscr{C}}$ is called the *family of coordinate functions*.

Typical choices of coordinates of interest can be statistical indicators $(\phi_U(V) \text{ can be a quantile of some observation of } V - U \text{ for instance})$. At a more abstract level, the triple $(\mathbf{S}, d_{\mathbf{S}}, (\phi_U)_{u \in \mathscr{C}})$ might as well encode a distance between realizations of random variables or between probability distributions.

Definition 2.9 (Model of margin - random case). A model of margin is the combination of a problem description, a family of probing sets and a set of coordinates of interest with the corresponding coordinate functions. It is denoted by:

$$\mathsf{M} = (\mathsf{E}, \mathscr{C}, \mathscr{A}, (\mathscr{G}_U)_{U \in \mathscr{C}}, \mathbf{S}, d_{\mathbf{S}}, (\phi_U)_{U \in \mathscr{C}}).$$
(8)

In the case where the coordinates of interest in \mathbf{S} are some deterministic "projections" of the random variables - i.e some statistical indicators - a first question is whether a model of margin in the random case can be replaced by a deterministic model of margins, reasoning only on these statistical indicators. In general, this is not the case, as it is useful to define the probing sets by imposing relationships between the laws of the random variables and not only on some of their statistical indicators. One can thus choose to explore some specific evolution of the random variables given by equations between them for instance.

Definition 2.10 (Margin - random case). Let $\mathsf{M} = (\mathsf{E}, \mathscr{C}, \mathscr{A}, (\mathscr{G}_U)_{U \in \mathscr{C}}, \mathbf{S}, d_{\mathbf{S}}, (\phi_U)_{U \in \mathscr{C}})$ be a given margin model and $U \in \mathscr{C}$ a variable of interest. The *margin* at the point U for the model M is defined as:

$$\mathbf{m}\left(U;\mathbf{M}\right) = \begin{cases} d_{\mathbf{S}}\left(\phi_{U}(U),\phi_{U}(\mathscr{A}^{c}\cap\mathscr{G}_{U|\mathscr{C}})\right) & \text{if } U \in \mathscr{A}, \\ -d_{\mathbf{S}}\left(\phi_{U}(U),\phi_{U}(\mathscr{A}^{c}\cap\mathscr{G}_{U|\mathscr{C}})^{c}\cap\phi_{U}(\mathscr{G}_{U|\mathscr{C}})\right) & \text{if } U \notin \mathscr{A}. \end{cases}$$
(9)

This definition is the same as in the deterministic case and is thus its generalization since the deterministic variables of interest can be interpreted as constant random variables. The spirit of this definition is that even if one deals with random variables, the margin should be ultimately expressed as a distance $d_{\mathbf{S}}$ in a deterministic metric set \mathbf{S} .

The acceptability property (Proposition 2.1) and the bounds on the margin increase (Proposition 2.2) also hold for the margin in the random case.

2.1.5 Example of the water heater in the random case

Let us go back to the example of Section 2.1.3. Let us model the variability of the heating time by making G a random variable - to take into account the fact that the outside temperature is uncertain and affects the heating time for instance. For given values of V and P, the heating time T is then also a random variable. The volume needed V_{needs} can also be modeled as a random variable, one realization representing the actual volume needed in one use.

The impacts on our model are the following. The state space remains the same $\mathbf{E} = \{ u = (P, V, T) \in \mathbb{R}^3 \}$, T still taking its values in \mathbb{R} . However, the fact that T is random has an effect on \mathscr{C} and \mathscr{A} . The set of problem constraints is then a subset of random variables

$$\mathscr{C} = \left\{ (P, V, T) \in \mathbb{R}^2 \left| \times \mathbb{L}^0 \left(\Omega, \mathbb{R} \right), 0 \le V \le V_{\max}, 0 \le P \le P_{\max} \text{ and } T = G \frac{V}{P} \right\},\right.$$

with G a given random variable.

The two constraints of the user must also be expressed in a random variable framework. Instead of the request $V > V_{\text{needs}}$, the user chooses $\mathbf{P}(V \leq V_{\text{needs}}) \leq 0.05$, which can be rewritten $V > q_{0.95}(V_{\text{needs}})$, where we recall that the quantity $q_{0.95}(V_{\text{needs}})$ is the quantile of V_{needs} of level 0.95. Instead of $T < T_{\text{lim}}$, they choose to focus on the average heating time $\mathbf{E}[T] < T_{\text{lim}}$. The acceptance set is then:

 $\mathscr{A} = \left\{ (P, V, T) \in \mathbb{R}^2 \times \mathbb{L}^0(\Omega, \mathbb{R}) | \mathbf{E}[T] < T_{\text{lim}} \text{ and } V > q_{0.95}(V_{\text{needs}}) \right\}.$ (10)

The probing sets are the same as in the deterministic case, except that T is considered now as a random variable - i.e an element of $\mathbb{L}^0(\Omega, \mathbf{E})$. In the case of Scenario 1, the projection of the non accepted set on the power P is now $\phi_U(\mathscr{A}^c) = (-\infty, \mathbf{E}[G] \frac{V}{T_{\text{lim}}}]$, the margin on P is then expressed $\mathbf{m}(U; \mathbf{M}_P) = P - \mathbf{E}[G] \frac{V}{T_{\text{lim}}}$. In order to define a margin on T, it is necessary to choose a deterministic coordinate of interest related to T. In this specific case of a condition on the average heating time, it is interesting to consider the expectation $\phi_u(P', V', T') = \mathbf{E}[T']$. In that case, the projection of the probed non acceptable set is $\phi(\mathscr{A}^c \cap \mathscr{G}_{u|\mathscr{C}}) = [T_{\text{lim}}; +\infty)$ and the margin on T is $T_{\text{lim}} - \mathbf{E}[T] = T_{\text{lim}} - \mathbf{E}[G] \frac{V}{P}$.

In Scenario 2, the margin on the projections on the power and on the volume have the same expressions as in the deterministic case, replacing V_{needs} by $q_{0.95}(V_{\text{needs}})$. For the heating time, it is possible to do the same reasoning as in the deterministic case, considering the expectation instead of a deterministic heating

time. The projection of the probed risky states is $\phi_u(\mathscr{A}^c \cap \mathscr{G}_{u|\mathscr{C}}) = \left(0, \frac{\mathbf{E}[G]V_{\text{needs}}}{\sqrt{V_{\text{needs}}^2 - V^2 + \left(\frac{\beta_V}{\beta_P}P\right)^2}}\right]$, which leads to the margin on the expected heating time $\frac{\mathbf{E}[G]V}{P} - \frac{\mathbf{E}[G]V_{\text{needs}}}{\sqrt{V_{\text{needs}}^2 - V^2 + \left(\frac{\beta_V}{\beta_P}P\right)^2}}$.

2.2 Models of margin in some specific cases

To conclude this section, let us highlight two types of margin models that encompass a wide variety of actual margin practices.

2.2.1 Directional margin

A model of *directional margin* consists in probing - i.e exploring - the states in one specific direction from the reference point and measuring the distance from this point to the unacceptable points in this specific direction. An intuitive interpretation of such a margin is that it answers the question "how far can the system go in this direction before reaching an unacceptable state?" Let the state space E be a vector space, the set of problem constraints \mathscr{C} and the acceptance set \mathscr{A} be two subsets of $\mathbb{L}^0(\Omega, \mathbb{E})$. For all $U \in \mathscr{C}$ and $w \in \mathbb{E} \setminus \{0\}$, consider the probing set

$$\mathscr{G}_U = \{U + \lambda w \,| \lambda \ge 0\},\$$

together with the family of coordinate functions $\phi_U: \mathscr{G}_U \to [0; +\infty)$ defined by:

$$\phi_U(U + \lambda w) = \lambda.$$

The associated margin m(U; M) is called the *directional margin* at U in the direction w. If $U \in \mathcal{A}$, it is written:

$$\mathsf{m}(U;\mathsf{M}) = \inf \left\{ \lambda \ge 0 \left| U + \lambda w \in \mathscr{A}^c \cap \mathscr{C} \right\} \right\}$$

Proposition 2.1 rewrites in this case:

Proposition 2.3 (Acceptability property - directional margins). For any $\lambda \geq 0$, if $\lambda < m(U; M)$ and $U + \lambda w \in \mathscr{C}$, then $U + \lambda w \in \mathscr{A}$.

If $U \notin \mathcal{A}$, the margin $\mathsf{m}(U; \mathsf{M})$ is nonpositive and its value is:

$$\mathbf{m}(U;\mathbf{M}) = \inf\left\{\lambda \ge 0 \left| U + \lambda w \in \mathscr{A}\right.\right\}.$$
(11)

This is proportional to the minimum distance to translate the state in the direction w before reaching the acceptance set.

Remark 2.1 (Discrete case). When E is discrete, it is possible to define a directional margin if E is a module on \mathbb{Z} and imposing $\lambda \in \mathbb{N}$.

2.2.2 Always acceptable box

A model of always acceptable box margin consists in finding a subset of the acceptance set \mathscr{A} that can be expressed as the Cartesian product of spaces of interest. An intuitive example of such a margin is the notion of *operating range* given on some relevant characteristics of a system - a temperature and a voltage range for instance. These ranges, given as descriptions of the system, imply that it operates correctly as long as each characteristic lies within its own range, the ranges being independent from the values of the other characteristics.

Let $(\mathbb{E}, \mathscr{C}, \mathscr{A})$ be a problem description, say in the random case; and $(\mathscr{G}_{U|\mathscr{C}})_{U\in\mathscr{C}}$ be a family of probing sets. Let **S** be the Cartesian product of metric spaces $\mathbf{S}_1, \ldots, \mathbf{S}_n$ with associated distances $d_{\mathbf{S}_i}, i \in [\![1; n]\!]$. Let $(\beta_1, \ldots, \beta_n)$ be a vector of positive numbers. β_i can be seen as the nominal value of coordinate \mathbf{S}_i . Then we can define the distance $d_{\mathbf{S}}$ on **S** by:

$$\forall s = (s_1, \dots, s_n), \quad \forall t = (t_1, \dots, t_n) \in \mathbf{S}, \quad d_{\mathbf{S}}(s, t) = \max_{1 \le i \le n} \frac{d_{\mathbf{S}}(s_i, t_i)}{\beta_i}.$$
 (12)

Let $U \in \mathscr{C}$, $U' \in \mathscr{G}_{U|\mathscr{C}}$, $\phi_U(U) = (s_1, \ldots, s_n)$ and $\phi_U(U') = (t_1, \ldots, t_n)$, Proposition 2.1 now writes: **Proposition 2.4** (Acceptability property - always acceptable box). If $d_{\mathbf{S}}(s_i, t_i) \leq \beta_i \mathbf{m}(U; \mathsf{M})$ for all $i \in [1; n]$, then $U' \in \mathscr{A} \cap \mathscr{C}$.

We can see that, in order for V to be acceptable, it is sufficient to check that each t_i is close enough to s_i independently from any cross effect with other variables (s_j, t_j) , with $i \neq j$.

When \mathbf{S}_i is a normed vector space, it is natural to define β_i as the norm of a nominal vector $\beta_i = ||x_i||$. An example in the setting of Section 2.1.3 is given in Figure 4. The point u = (P, V, T) is projected on the plane $(P, V) \in \mathbb{R} \times \mathbb{R}$. The coefficients β_1, β_2 are given by nominal values β_P, β_V . The margin **m** is visualized as a function of the size of rectangle (P, V), namely its width and its length divided by the nominal values. For all (P', V', T') in the probing set, the acceptability property is, for U = (P, V, T):

If
$$|P' - P| < \mathsf{m}(U; \mathsf{M}) \beta_P$$
 and $|V' - V| < \mathsf{m}(U; \mathsf{M}) \beta_V$, then $(P', V', T') \in \mathscr{A}$



Figure 4: Example of an always acceptable box in the example of the water heater

3 Defining the acceptance set: general risk measures

3.1 General risk measures

The acceptance set \mathscr{A} , defined in Sections 2.1 and 2.6, is a formal way to represent how the risk is conceived by the person in charge of the margin definition. Providing a rigorous relevant model to measure how the risk is perceived by different stakeholders has been and is still an important topic in risk management. It is the core of decision theory, in which this problem is considered with the angle of modeling the behavior of economic agents. Von Neuman and Morgenstern [31] proposed a formalism with the concept of expected utility in 1953.

The quantification of risk also regained some interest in finance, in the early 2000's with an increasing need of regulation practices. In this context, Artzner, Delbaen, Eber and Heath introduced the concept of coherent risk measures, explicitly linked with the expected utility [9, ch 2 and 4]. Their use in other engineering fields followed [11, 26], especially in robust optimization under uncertainty [25], thanks to their min-max formulation.

These uses in the literature motivate the introduction of the notion of the general risk measure:

Definition 3.1 (General risk measure). Let E be a state space and let (\mathscr{R}, \preceq) be a partially ordered set. In the deterministic case, a \mathscr{R} -valued risk measure is a function from E to \mathscr{R} . In the random case, a \mathscr{R} -valued risk measure is a function from $\mathbb{L}^0(\Omega, E)$ to \mathscr{R} .

In practice, \mathscr{R} will often be the real line or \mathbb{R}^d .

The risk measure can take a lot of different forms and depends on the needs of a specific application. Immediate examples in the case of random variables of interest are the expectation, the variance, the probability to exceed a threshold. In most engineering practices, risk measures are used but are implicit. We found that in a lot of cases in engineering under uncertainty, it is possible to exhibit a risk measure that creates an acceptance set and can be interpreted as a quantification of risk.

In the sequel, our convention is that the greater the risk measure, the riskier the situation. In other terms $\rho(U_2) \leq \rho(U_1)$ implies that the variable of interest U_1 is considered more risky than U_2 . The opposite convention can be found in the literature and in practices (as it is the case with the Capability process in Section 4.1), with quantities measuring the safety or the performance of the variable of interest. They can be converted to risk measures in our sense with a change of sign.

Once a risk measure ρ is fixed, an acceptable risk value ρ_{req} is chosen, taking into account the context of the study. The acceptance set is then the variables of interest with a risk below the threshold:

Definition 3.2 (Acceptance set associated to a risk measure). Let E be a state space, let ρ be a \mathscr{R} -valued risk measure and ρ_{req} an element of \mathscr{R} . The acceptance set associated to the risk measure ρ with a risk level ρ_{req} is defined in the deterministic case by:

$$\mathscr{A} = \{ u \in \mathcal{E} \mid \rho(u) \preceq \rho_{\operatorname{req}} \}.$$

In the random case, its expression is:

$$\mathscr{A} = \left\{ U \in \mathbb{L}^0 \left(\Omega, \mathbf{E} \right) | \rho(U) \preceq \rho_{\mathrm{req}} \right\}$$

The monetary risk measures and coherent risk measures from [1] are of course in the family of the general risk measures. They are supported by a valuable framework of mathematical tools. However, it appears that these definitions are too restrictive to describe the wide variety of engineering practices and thus, in our descriptive approach, we cannot restrict the study to the sole monetary measures. For instance, the variance is not a monetary risk measure, but it is important in an engineering context to be able to address such a general risk measure.

Definition 3.3 (Margin on the risk level). Let $\mathsf{M} = (\mathsf{E}, \mathscr{C}, \mathscr{A}, (\mathscr{G}_U)_{U \in \mathscr{C}}, \mathbf{S}, d_{\mathbf{S}}, (\phi_U)_{U \in \mathscr{C}})$ be a margin model in which the acceptance set if defined by a risk measure $\rho : \mathbb{L}^0(\Omega, \mathsf{E}) \to (\mathscr{R}, \preceq)$ and a risk level ρ_{req} through Definition 3.2. The margin on the risk level is defined by

$$\mathsf{m}_{\mathrm{risk}}(U;\mathsf{M}) = \rho_{\mathrm{req}} - \rho(U). \tag{13}$$

Remark 3.1. The general risk measures are introduced because they often have a meaning in the considered field and enable to define an acceptance set in an intuitive manner. It is however always possible to define a risk measure given an arbitrary set \mathscr{A} , using the indicator function $\rho(I) = 1 - \mathbf{1}_{\mathscr{A}}(I)$ and any threshold ρ_{req} in (0, 1).

Remark 3.2. (Composition of margins) Let us remark that if m(U; M) is positive, then U is acceptable and if m(U; M) is negative, then U is not acceptable. A margin, associated with a margin model M_1 , can thus be used as the opposite of a risk measure $\rho_1(U) = -m(U; M_1)$. The risk threshold can be interpreted as the opposite of a demanded margin $\rho_{req} = -m_{demanded}$. The related acceptance set $\mathscr{A} = \{X | \rho_1(X) \leq \rho_{req}\}$ can be used to define another margin model M_2 (such as in Section 5.2). The interpretation of the margin model M_2 is the measure of the extra quantity that is left once a margin $m_{demanded}$ has been imposed in M_1 .

The expression of the risk measure depends inherently on the cases (see Section 4 for applied examples). Its choice takes into account the models used, the criticity of the quantities of interest and the purpose of the modeling.

Remark 3.3. In the unidimensional case $E \subset \mathbb{R}$ and $\mathscr{R} = \mathbb{R}$, if the risk measure ρ is additive in the sense that $\rho(u+m) = \rho(u) + m$ for any $m \in \mathbb{R}$, which is related to the notion of translation invariance for monetary risk measures [9], the directional margin in the sense of Section 2.2.1, in the direction w = +1 is expressed:

$$m(u; \mathsf{M})) = \inf \left\{ \lambda \in \mathbb{R}^+ | \rho(u + \lambda) > \rho_{\mathrm{req}} \right\}$$

= $\inf \left\{ \lambda \in \mathbb{R}^+ | \lambda > \rho_{\mathrm{req}} - \rho(u) \right\}$
= $\rho_{\mathrm{req}} - \rho(u).$ (14)

Thus the margin at u coincide with the margin on the risk level.

4 Construction of models of margin for classical margins

The previous sections were dedicated to the construction a formal framework, in which a *model of margin* was defined (Definition 2.4) and a way to compute a *margin* at a point for a model of margin was given (Definition 2.5).

In this section, we illustrate this formalism by interpreting the margins from the literature within our framework - and then demonstrate the possibility to recast some concepts of margins with models of margin.

4.1 The Process Capability Index

The Process Capability Index or Process Capability Ratio C_p is used in the field of quality control with statistical methods and it is a part of the six-sigma method [19, 28, ch 12]. The idea is to link a probability of exceedance of a normal random variable to a ratio between its expectation and its standard deviation. This ratio is the capability process. We choose to focus on one particular expression, but a similar reasoning can be done with other indicators defined in [24, sec 8.3] for instance. The phenomenon is described by a real normal random variable X. The state space is then $\mathbf{E} = \mathbb{R}$ and the set of problem constraints is the set of normal random variables $\mathscr{C} = \{X \sim \mathcal{N}(\mu, \sigma^2) | (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}^+_*\}$. The goal of the C_p is to measure the probability that $X \in \mathscr{C}$ remains in the interval bounded by the lower specification limit $\mathrm{LSL} \in \mathbb{R}$ and the upper specification limit $\mathrm{USL} \in \mathbb{R}$. The expression of the capability process is the following:

$$C_p(X) = \min\left(\frac{\mu - \text{LSL}}{3\sigma}, \frac{\text{USL} - \mu}{3\sigma}\right)$$
(15)

with $\mu = \mathbf{E}[X]$ and $\sigma = \sqrt{\operatorname{Var}(X)}$.

As we will show, this indicator bounds the probability of exceedance by lower and upper limit with a value $r \in [0, 1]$. The acceptance set of the related model of margin is then:

$$\mathscr{A} = \left\{ X \in \mathbb{L}^0 \left(\Omega, \mathcal{E} \right) | \mathbf{P} \left(X > \mathcal{USL} \right) < r \text{ and } \mathbf{P} \left(X < \mathcal{LSL} \right) < r \right\}.$$
(16)

As described in Section 3.1, this acceptance set is defined thanks to an associated risk measure and threshold:

$$\rho(X) = \max\left(\mathbf{P}\left(X > \text{USL}\right), \mathbf{P}\left(X < \text{LSL}\right)\right) < \rho_{\text{req}} = r.$$
(17)

We are interested in all possible changes of the random variable X, and then $\mathscr{G}_X = \mathbb{L}^0(\Omega, \mathbb{R})$ and $\mathscr{G}_{X|\mathscr{C}} = \{Y \sim \mathscr{N}(\mu, \sigma^2) | (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}^+_*\}$. The coordinate of interest is a real deterministic variable, i.e $\mathbf{S} = \mathbb{R}$ and has an expression similar to the capability process:

$$\phi_X(Y) = \phi(Y) = \min\left(\frac{\text{USL} - \mathbf{E}[Y]}{\varphi_{1-r}\sqrt{\text{Var}(Y)}}, \frac{\mathbf{E}[Y] - \text{LSL}}{\varphi_{1-r}\sqrt{\text{Var}(Y)}}\right)$$

with φ_{1-r} being the quantile of level 1-r of a standard normal random variable. We denote by $C_{p,r}(Y)$ the right-hand side. With this expression, we can prove that the projection of the non acceptable coordinates is:

$$\phi_X\left(\mathscr{A}^c \cap \mathscr{G}_{X|\mathscr{C}}\right) = (-\infty; 1], \tag{18}$$

and thus $\phi_X (\mathscr{A}^c \cap \mathscr{G}_X)^c \cap \phi_X (\mathscr{G}_{X|\mathscr{C}}) = (1; +\infty)$. By Definition 2.10, assuming that $X \in \mathscr{A} \cap \mathscr{C}$, the margin at X is:

$$m(X; M) = d_{\mathbb{R}}(C_{p,r}(X), (-\infty; 1]) = C_{p,r}(X) - 1$$

A choice of r = 0.001 leads to $\varphi_{1-r} \simeq 3$ and thus justifies the expression of $C_p \simeq C_{p,0.001}$.

The quantity $C_p - 1$ can then be directly interpreted as a margin with the acceptance set given in Equation (16). This margin is consistent with the intuition: if C_p is greater than 1, the condition expressed in Equation (16) is satisfied and the margin is positive. If C_p is lower than 1, the condition (16) is not satisfied and the margin is negative. Constructing this model of margin has also enabled to exhibit a family of capability processes $C_{p,r}$ associated to a risk level $\rho_{req} = r$.

Proof of Equation (18) Let us prove that $\phi_X \left(\mathscr{A}^c \cap \mathscr{C}_X \right) = (-\infty; 1]$. First, we can rewrite the upper exceedance probability of the acceptance set as $\mathbf{P}(X > \mathrm{USL}) = \mathbf{P}\left(\frac{X - \mathbf{E}[X]}{\sqrt{\mathrm{Var}(X)}} - \frac{\mathrm{USL} - \mathbf{E}[X]}{\sqrt{\mathrm{Var}(X)}} > 0\right)$. X being a normal random variable, $\frac{X - \mathbf{E}[X]}{\sqrt{\mathrm{Var}(X)}}$ is a standard normal. Thus, $\mathbf{P}\left(\frac{X - \mathbf{E}[X]}{\sqrt{\mathrm{Var}(X)}} > \frac{\mathrm{USL} - \mathbf{E}[X]}{\sqrt{\mathrm{Var}(X)}}\right) < r$ is equivalent to $\frac{\mathrm{USL} - \mathbf{E}[X]}{\sqrt{\mathrm{Var}(X)}}$ being greater than φ_{1-r} . In other terms:

$$\mathbf{P}\left(X > \mathrm{USL}\right) < r \Leftrightarrow \frac{\mathrm{USL} - \mathbf{E}\left[X\right]}{\varphi_{1-r}\sqrt{\mathrm{Var}\left(X\right)}} > 1.$$

The same can be done with LSL, and we finally have:

$$X \in \mathscr{A}^{c} \cap \mathscr{G}_{X|\mathscr{C}} \Leftrightarrow \min\left(\frac{\mathrm{USL} - \mathbf{E}[X]}{\varphi_{1-r}\sqrt{\mathrm{Var}(X)}}, \frac{\mathbf{E}\left[X\sqrt{\mathrm{Var}(X)}\right] - \mathrm{LSL}}{\varphi_{1-r}\sqrt{\mathrm{Var}(X)}}\right) \leq 1.$$

4.2 Reliability indices

Consider a system described by a pair of random variables R and S, respectively denoting the *resistance* of the system and the *stress* it undergoes, so that the system is in failure whenever R < S. Letting $\mu = \mathbf{E} [R - S]$ and $\sigma^2 = \text{Var} (R - S)$, we define the *Cornell reliability index* [2, 20] by

$$\beta_{\rm Co} = \frac{\mu}{\sigma}.$$

The practical interest of this index is that, if R - S is a Gaussian random variable, then the *probability* of failure $\mathbf{P}(R < S)$ is given by $\Phi(-\beta_{Co})$, where Φ is the cumulative distribution function of the standard Gaussian distribution. Therefore, this probability, which may be expected to be very small and therefore difficult to estimate accurately, can actually be computed from the estimation of the first two moments of the random variable R - S.

From its definition, one may directly interpret the Cornell index as a *distance* between μ and 0, normalized by the standard deviation σ . This suggests that β_{Co} may be recasted as a margin, and we now construct a model of margin allowing this interpretation.

Since the system is described by the pair U = (R, S), we take as state space $E = \mathbb{R}^2$. The set of problem constraints \mathscr{C} is the set of all pairs of random variables (R, S) in \mathbb{R}^2 , such that $\operatorname{Var}(R - S) < \infty$, and the acceptance set is defined by

$$\mathscr{A} = \{ (R, S) \in \mathbb{L}^0 \left(\Omega, \mathbb{R}^2 \right) | \mathbf{E} [R - S] \ge 0 \}.$$

For any $U \in \mathscr{C}$, we set $\mathscr{G}_U = \mathscr{C}$ and define, for any $U' = (R', S') \in \mathscr{C}$,

$$\phi_U(U') = \frac{\mathbf{E}\left[R' - S'\right]}{\sqrt{\operatorname{Var}\left(R' - S'\right)}} \in \mathbf{S} = \mathbb{R} \cup \{-\infty, +\infty\}.$$

With this model of margin M at hand, we may now remark that $\phi_U(\mathscr{A}^c \cap \mathscr{G}_{U|\mathscr{C}}) = [-\infty, 0)$, so that for any $U \in \mathscr{C}$, the margin at U satisfies

$$\mathsf{m}(U;\mathsf{M}) = \beta_{\mathrm{Co}}.$$

As a consequence, our framework provides a natural formalization of the Cornell index as a margin, with respect to the margin model M.

The margin model M constructed above is an instance of the random case described in Definition 2.9, and this is quite natural since the system is described by a pair of random variables. However, up to introducing a parametrization of the law of the random variable Z = R - S, it is just as possible to interpret the Cornell index within a *deterministic* margin model. Assume indeed that the system is now described by the pair $u = (\mu, \sigma)$ taking its values in the state space $E = \mathbb{R} \times [0, +\infty)$, and let

$$\mathscr{C} = \mathcal{E}, \qquad \mathscr{A} = \{ u = (\mu, \sigma) \in \mathcal{E} : \mu \ge 0 \}.$$

For any $u \in \mathscr{C}$, take $\mathscr{G}_u = \mathscr{C}$ and define, for any $u' = (\mu', \sigma') \in \mathscr{C}$,

$$\phi_u(u') = \frac{\mu'}{\sigma'} \in \mathbf{S} = \mathbb{R} \cup \{-\infty, +\infty\}$$

Then again, the margin at $u \in \mathscr{C}$ associated to this margin model M writes $\mathsf{m}(u;\mathsf{M}) = \beta_{\mathrm{Co}}$.

A similar parametrization allows to express another common reliability indicator, the *Hasofer-Lind index*, as a margin within a (deterministic) margin model. We now assume that the system is described by

a random vector $X = (X_1, \ldots, X_n) \in \mathbb{R}^n$ and a function $G : \mathbb{R}^n \to \mathbb{R}$, with the property that the system is in failure if and only if G(X) < 0. Assume moreover that there exists a function $T : \mathbb{R}^m \to \mathbb{R}^n$, called an *isoprobabilistic transform*, such that if U is a standard Gaussian random vector in \mathbb{R}^m , which we denote by $U \sim \mathcal{N}_m(0_{\mathbb{R}^m}, I_{m \times m})$, then T(U) has the same law as X. We call \mathbb{R}^m the *standardized space*, and define

$$F = \{ u \in \mathbb{R}^m | G(T(u)) < 0 \},\$$

as well as the design point

$$u^* = \underset{u \in F}{\operatorname{arg\,min}} \|u\|.$$

The Hasofer-Lind index is defined by

 $\beta_{\mathrm{HL}} = \|u\|.$

In any reasonable model, the failure event G(X) < 0 should be rare, so that F should be located far from the origin of the standardized space, and therefore β_{HL} should be positive - and even large.

The so-called FORM approximation [20] relates the Hasofer-Lind with the probability of failure by the formula

$$\mathbf{P}\left(G(X) < 0\right) \simeq \Phi(-\beta_{\mathrm{HL}}),$$

which indicates a similarity between the Cornell and Hasofer-Lind indices. To pursue this analogy, we now construct a margin model based on the isoprobabilistic transform T. We take as a state space E the standardized space \mathbb{R}^m , let $\mathscr{C} = E$, $\mathscr{A} = \{u \in \mathbb{R}^m | u \notin F\}$, $\mathscr{G}_u = \mathscr{C}$ for any $u \in \mathscr{C}$, and $\phi_u(u') = u' \in \mathbf{S} = \mathbb{R}^m$. Then, with this margin model M and under the assumption that $0_{\mathbb{R}^m} \notin F$, which is equivalent to $\beta_{\text{HL}} > 0$, we may conclude that the Hasofer-Lind index rewrites

$$\beta_{\mathrm{HL}} = \mathsf{m}\left(0_{\mathbb{R}^m};\mathsf{M}\right),\,$$

so that it is the margin at the origin of the standardized space.

4.3 Safety coefficients

In order to introduce the safety coefficients, we are considering a generalization of the resistance-stress case presented in Section 4.2. Let us consider a phenomenon modeled by n deterministic real variables (x_1, \ldots, x_n) , with a failure condition of the form:

The state is in the failure domain if and only if
$$G(x_1, \ldots, x_n) < 0$$
, (19)

with G being interpreted as a measure of the safety - the greater it is, the safer is the state according to the model. Furthermore, let us assume that for all i in [1; n], the variable x_i is nonegative and the function $x_i \mapsto G(x_1, \ldots, x_i, \ldots, x_n)$ is nonincreasing, the other variables being fixed. The x_i can be seen as some stress variables. In this context, the safety coefficients are defined as a tuple:

$$(\gamma_1, \dots, \gamma_n) \in [1; +\infty)^n \tag{20}$$

such that a configuration $(x_1, \ldots, x_n) \in \mathbb{R}^n_+$ is not satisfactory as soon as $G(\gamma_1 x_1, \ldots, \gamma_n x_n) < 0$. We want that a state in the failure domain to not be satisfactory; this motivates the condition that the value of the safety coefficients are greater than or equal to 1.

The safety coefficients can be interpreted as a stricter condition imposed on the system so as to cover some uncertainties and consequences that has not been modeled by G and the inputs x_1, \ldots, x_n . This stricter condition takes the form of an increase in the variables having a negative impact on the safety function. It can thus be interpreted as a margin and we will derive the associated model of margin.

The state space is composed of the variables describing the phenomenon, namely $E = \{(x_1, \ldots, x_n) \in \mathbb{R}^n_+\}$. In our case we have no equations, although there might be some, so the variables can be independently chosen one from the others, giving the set of problem constraints $\mathscr{C} = E$. The acceptance set is given by the failure function:

$$\mathscr{A} = \{u = (x_1, \dots, x_n) | G(u) \ge 0\}.$$

For a given vector of safety coefficients $\gamma = (\gamma_1, \ldots, \gamma_n) \in [1; +\infty)^n$, we can also define a satisfactory set for γ as:

$$\mathscr{A}_{\gamma} = \left\{ u = (x_1, \cdots, x_n) | G(\gamma_1 x_1, \dots, \gamma_n x_n) \ge 0 \right\}.$$
(21)

The condition that an unacceptable state must not be satisfactory can be rewritten $\mathscr{A}_{\gamma} \subset \mathscr{A}$. We want a more precise expression, of the form

$$u \in \mathscr{A}_{\gamma} \Leftrightarrow \mathsf{m}(u;\mathsf{M}) \ge m_{\gamma} \ge 0$$

for a model of margin M. Thanks to the monotonicity of G, we can focus only on the vectors which components are each greater than or equal to (x_1, \ldots, x_n) . The probing set at $u = (x_1, \ldots, x_n)$ is:

$$\mathscr{G}_u = \{(y_1,\ldots,y_n) \in \mathbb{R}^*_+ | x_1 \leq y_1,\ldots,x_n \leq y_n\}.$$

The coordinates of interest \mathbf{S} are the relative changes of the variable, i.e.

$$\phi_u(y_1,\ldots,y_n) = \left(\frac{y_1}{x_1},\ldots,\frac{y_n}{x_n}\right)$$

For $\gamma \in [1; +\infty)^n$, $u = (x_1, \ldots, x_n)$ in \mathscr{C} , we have:

u

$$\in \mathscr{A}_{\gamma} \Leftrightarrow \forall v \in \mathscr{G}_{u|\mathscr{C}}, \quad \text{if} \quad \forall i \in [\![1;n]\!], \quad \frac{y_i}{x_i} < \gamma_i, \quad \text{then} \quad v \in \mathscr{A} \\ \Leftrightarrow \forall v \in \mathscr{G}_{u|\mathscr{C}}, \quad \text{if} \quad \forall i \in [\![1;n]\!], \quad |\phi_u(v)_i - \phi_u(u)_i| < \gamma_i - 1 \quad \text{then} \quad v \in \mathscr{A} \\ \Leftrightarrow \forall v \in \mathscr{G}_{u|\mathscr{C}}, \quad \text{if} \quad \max_{i \in [\![1;n]\!]} \frac{|\phi_u(v)_i - \phi_u(u)_i|}{\beta_i} < \sum_{j=1}^n (\gamma_j - 1) \quad \text{then} \quad v \in \mathscr{A},$$

$$(22)$$

where $\beta_i = \frac{\gamma_i - 1}{\sum_{j=1}^n (\gamma_j - 1)}$. The last term of expression (22) is in fact the acceptability property of an always acceptable box margin, defined in Section 2.2.2. This provides the distance $d_{\mathbf{S}}(s, t) = \max_{1 \le i \le n} \frac{|s_i - t_i|}{\beta_i}$, to complete our model of margin $\mathsf{M} = (\mathsf{E}, \mathscr{C}, \mathscr{A}, (\mathscr{G}_u)_{u \in \mathscr{C}}, \mathbf{S}, d_{\mathbf{S}}, (\phi_u)_{u \in \mathscr{C}})$. We can state our final result:

$$u \in \mathscr{A}_{\gamma} \Leftrightarrow \mathsf{m}(u;\mathsf{M}) \ge m_{\gamma} = \sum_{j=1}^{n} (\gamma_j - 1).$$
 (23)

The safety coefficient condition, given by \mathscr{A}_{γ} in Equation (21), is reintepreted as an always acceptable box margin on the relative variations of the variables.

The margin model M is not completely independent of the choice of γ , but is a function of the relative weights of its components, namely the $\beta_i = \frac{\gamma_i - 1}{\sum_{j=1}^{n} (\gamma_j - 1)}$. Multiplying the demanded margin m_{γ} by $\lambda > 0$ would then be equivalent to imposing new coefficients of value $\gamma_i^* = 1 + \lambda(\gamma_i - 1)$. In particular, we can only compare the margins for sets of coefficients with the same relative weights.

The monotonicity assumption is quite strong but it is satisfied for example in the linear elasticity domain in Civil Engineering. The works described in Section 4.6.3 show an example of such assumptions with references to variables of type "the lower the better" or "the greater the better".

Margin of safety Let us apply the method of safety coefficients to the resistance-stress approach described in Section 4.2. We consider a system described by a (deterministic) pair $(r, s) \in \mathbb{R}^2_+$, which is in failure if and only if r < s. Then one may prescribe two safety coefficients $\gamma_s > 1 > \gamma_\tau$ and consider as satisfactory any configuration (r, s) satisfying the condition $\gamma_\tau r > \gamma_s s$. For such a configuration, the *Margin of Safety* is defined in the aeronautic literature [15, sec 3] [16, sec 3.6] as $MOS = \frac{\gamma_\tau r}{\gamma_s s} - 1$. It measures in a certain sense the remaining margin, once the safety coefficients have been imposed.

4.4 Phase margin and gain margin

In the theory of control and stability of linear systems, the gain margin and the phase margin are common indicators to demonstrate the stability or the instability of a system [10, Sec 6.4],[21, Sec 8.4.]. It is possible to build models of margin describing the two types of margins.

Let us first recall the definition of phase margin and the gain margin. For a simple feedback system of one variable, the definition of the BIBO (Bounded Inputs leads to Bounded Outputs) stability is that the poles of its closed-loop transfer function $H_{cl}(s)$ must all have a negative real part. A transfer function is a complex rational function (i.e it belongs to $\mathbb{C}(X)$). The closed-looped transfer function can be expressed as $H_{cl}(s) = \frac{H_{fo}(s)}{1+H_{ol}(s)}$, with $H_{ol}(s)$ the open-loop transfer function and $H_{fo}(s)$ the forward transfer function. The poles of H_{cl} are given by the (complex) roots of $1 + H_{ol}(s)$ and the poles of $H_{fo}(s)$.

The Nyquist stability test [21, Sec 8.2] gives a graphical condition to count the complex roots of the function $1 + H_{\rm ol}(s)$. Under some quite general assumptions on the transfer function, this test is sufficient to prove the BIBO stability. This test is motivated by Cauchy's argument principle: the number of roots of $1 + H_{\rm ol}(s)$ is given by the number of clockwise encirclements of the origin by the curve in the complex plane $1 + H_{\rm ol}(j\omega)$ when ω goes from $-\infty$ to $+\infty$, j being the imaginary unit. This test is equivalently reformulated by counting the number of clockwise encirclements of the point -1 by $H_{\rm ol}(j\omega)$. In order to characterize if the curve encircles the point -1, two criteria have been exhibited.

Gain Margin The first criterion focuses on the lowest abscissa of the points at the intersection between $(H_{\rm ol}(j\omega))_{\omega \in \mathbb{R}}$ and the real axis. If this abscissa is greater than -1 then the curve $(H_{\rm ol}(j\omega))_{\omega \in \mathbb{R}}$ will only pass "to the right" of -1 and thus this point is likely not to be encircled (see Figure 5). This motivates the definition of the *gain margin* as the logarithm of the ratio between -1 and the lowest abscissa of the intersecting points with the real axis, when this abscissa is negative. If we denote by $\omega_{\rm PC}$ the (possibly infinite) frequency at which the infimum is reached, then the gain margin is:

$$GM = 20 \log \left(\frac{-1}{H_{\rm ol}(j\omega_{\rm PC})}\right) = -20 \log(|H_{\rm ol}(j\omega_{\rm PC})|).$$

$$(24)$$

It is expressed in decibel. Figure 5 shows the plot of two open loop functions $H^1_{\rm ol}(j\omega)$ and $H^2_{\rm ol}(j\omega)$ in the complex plane (also known as the Nyquist plots). The blue one corresponds to an unstable system as the -1 point is encircled and consequently the margin is negative. The orange one corresponds to a stable system as there is no encirclement and thus the margin is positive.

Phase Margin The second criterion focuses on the lowest of the arguments of the points at the intersection between $(H_{\rm ol}(j\omega))_{\omega\in\mathbb{R}}$ and the unit circle $|s| = 1, s \in \mathbb{C}$ - the arguments being taken between 0 and 2π rad. If this argument is greater than π , then the point -1 is likely not to be encircled, using the symmetry and the fact that $H_{\rm ol}(j\omega)$ vanishes at infinity. If we denote by $\omega_{\rm GC}$ the frequency at which the infimum is reached, the phase margin is expressed as:

$$PM = Arg(H(j\omega_{GC})) - \pi, \qquad (25)$$

with $\operatorname{Arg}(z)$ returning the argument of z in $[0, 2\pi)$.

In Figure 6, two transfer functions are plotted as well as the unit circle. We can see that the blue one that corresponds to an unstable system also has a negative Phase Margin. On the contrary, the orange one that corresponds to a stable system has a positive margin.

It must be noted that there might be some cases in which these margins are in contradiction with the intuition (i.e having a negative margin for a stable system [21, Sec 8.4.2]) and a consistency check with the Nyquist plot is required. In the following, we assume that we are working on systems in which stability is equivalent to positive margins.

Model of margin A way to build our gain and phase margin models is to consider that the state space is the set of the transfer functions, represented by the rational functions $E = \mathbb{C}(X)$, the set of the problem constraints is the subset of the systems for which the margins are relevant $\mathscr{C} \subset E$. The acceptance set is the transfer functions of stable systems according to the Nyquist test:

$$\mathcal{A} = \left\{ \left. H_{\rm cl}(s) = \frac{H_{\rm fo}(s)}{1 + H_{\rm ol}(s)} \in \mathbf{E} \right| 1 + H_{\rm ol}(s) = 0 \Rightarrow \Re(s) < 0 \right\}.$$
(26)

We investigate all the possible evolutions and thus $\mathscr{G}_{u|\mathscr{C}} = \mathscr{C}$.

Gain margin In the case of the gain margin, the coordinate of interest is the level (in dB) of the lowest abscissa of the intersection between the real axis and the curve given by the open loop.

$$\phi_u(H_{\rm cl}) = 20\log(A(H_{\rm cl})) \tag{27}$$

where $A(H_{cl}) = \inf \{H_{ol}(j\omega) \text{ s.t } \Im(H_{ol}(j\omega)) = 0, \omega \in \mathbb{R}\}$. The coordinate space and the distance are simply $(\mathbf{S}, d_{\mathbf{S}}) = (\mathbb{R}, |\cdot|)$. An unstable system leads to a level lower than 0 dB and then $\phi(\mathscr{A}^c \cap \mathscr{C}) = (-\infty, 0]$. If we denote by ω_{PC} the (possibly infinite) frequency at which the infimum in the definition of A is obtained, the value of the margin is:

$$\mathsf{m}(H_{\rm cl};\mathsf{M}_{\rm GM}) = -20\log(|H_{\rm ol}(j\omega_{\rm PC})|) = \mathrm{GM}$$

Phase margin In the case of the phase margin, the coordinate of interest is the lowest argument of the points at the intersection between the unit circle and the open loop:

$$\phi_u(H_{\rm cl}) = \inf \left\{ \operatorname{Arg}(H_{\rm ol}(j\omega)) | H_{\rm ol}(j\omega) | = 1, \omega \in \mathbb{R} \right\},\tag{28}$$

with $\operatorname{Arg}(z)$ in $[0, 2\pi)$. The coordinate space and the distance are $(\mathbf{S}, d_{\mathbf{S}}) = (\mathbb{R}, |\cdot|)$. An unstable system leads to an argument lower than π and then $\phi(\mathscr{A}^c \cap \mathscr{C}) = (0, \pi]$. If we denote by ω_{GC} the (possibly infinite) frequency at which the infimum of Equation (28) is obtained, the value of the margin is:

$$\mathbf{m}(H_{\rm cl}; \mathbf{M}_{\rm PM}) = \operatorname{Arg}(H_{\rm ol}(j\omega_{\rm GC})) - \pi = \mathrm{PM}.$$
(29)



Figure 5: Gain Margin (GM) for the transfer functions with open loops $H_{\rm ol}^1$ and $H_{\rm ol}^2$, represented in the complex plane.

4.5 Monetary risk measures

The coherent risk measures used in finance and robust probabilistic optimization are a special case of our definition of general risk measures [25, 1]. As they focus on a single quantity $X \in \mathbb{R}$ (the monetary loss in finance, an aggregated cost in optimization) they offer some good properties such as convexity.

Let us denote by ρ a monetary risk measure and construct a related margin model. The state space is the space $E = \mathbb{R}$ in which this one dimensional quantity takes its values. The problem constraints is the space of random variables on E, $\mathscr{C} = \mathbb{L}^0(\Omega, E)$. A convention is that the threshold of a monetary risk measure is 0, so that the acceptance set is:

$$\mathscr{A} = \left\{ X \in \mathbb{L}^0 \left(\Omega, \mathbb{R} \right) | \rho(X) \le 0 \right\}.$$

We choose a directional model of margin M, defined in Section 2.2.1, in the direction of an increase in X. The expression of the margin is then:

$$\mathbf{m}\left(X;\mathbf{M}\right) = d_{\mathbf{S}}(0, \{\lambda \in \mathbb{R}_+ | \rho(X+\lambda) > 0\}). \tag{30}$$



Figure 6: Phase Margin (PM) for the transfer functions with open loops $H_{\rm ol}^2$ and $H_{\rm ol}^3$, represented in the complex plane.

A property of the monetary margin is the translation invariance, i.e $\rho(X + \lambda) = \rho(X) + \lambda$. Thus:

$$\mathbf{m}(X;\mathbf{M}) = d_{\mathbf{S}}(0, \{\lambda \in \mathbb{R}_+ | \lambda > -\rho(X)\}) = -\rho(X).$$
(31)

We can then interpret the risk measure as the opposite of the margin in the direction of an increase in X. As X represents a cost or a loss, the margin can be seen as the maximum cost/loss increase before not being acceptable. Interestingly, the margin on the risk level (Section 3.3) associated to this model of margin is also $0 - \rho(X) = -\rho(X)$. The monetary risk measure has the double interpretation of being both a margin on the maximum cost/loss increase and a margin on the risk level.

4.6 Existing frameworks for margins

Some works focused on developing frameworks and guidelines to take into account uncertainties using margins. This section aims at giving an interpretation of their approach through the prism of the margin model.

4.6.1 Task group on safety margins (SMAP) and Risk-Informed Safety Margins Characterization

In Civil Nuclear Safety, the Nuclear Energy Agency (NEA) set up a task group on safety margins (SMAP). Its purpose was to propose a framework for a regulator to spot a change in the safety margins when the design of a nuclear power plant is modified with respect to an original design [17]. To do so, they proposed a margin definition, that was then adopted in the nuclear industry in the Risk-Informed Safety Margins Characterization (RISMC) approach [16, 27][5, pp 44].

Two types of margins are exhibited in these works. The first one is the Cornell index, defined in the resistance-stress case in Section 4.2, with R and S independent variables. It is called margin to damage

and its expression is:

$$MD = \frac{\mathbf{E}[R] - \mathbf{E}[S]}{\sqrt{Var(R) + Var(S)}}.$$
(32)

In these works, the authors highlighted the fact that such a margin is not informative enough and is not sufficient in the case of nuclear safety - it makes a Gaussian assumption on R and S in order to link it with a probability of failure, for instance. They prescribed the use of another margin, namely the real failure probability $\mathbf{P}(S > R)$, that is interpreted in our framework as a margin on a risk measure/statistical indicator. It is interesting to notice that they do not define some threshold on the probability. This is explained because these thresholds should have already been fixed during the initial design. Their decision criteria are based on the relative changes of margins - and thus changes of $\mathbf{P}(S > R)$ - instead of their actual values.

4.6.2 Quantification of margin and uncertainties (QMU)

Another concept of margin is used in the context of Nuclear Safety, in the field of "Quantification of margins and uncertainties" (QMU) that defines itself as a subtopic of uncertainty quantification and sensitivity analysis [13]. The specificity of the approach is the focus on the computation of margins while keeping two distinct sources of uncertainty [13]:

- the epistemic uncertainty "derives from a lack of knowledge about the appropriate value to use for a quantity that is assumed to have a fixed value in the context of a particular analysis" ([13, sec 2]). It can be modeled by a random vector $e \in \mathbb{L}^0(\Omega_e, \mathbb{R}^{d_e})$ although other representations are possible [14].
- the aleatory uncertainty arises from an inherent randomness in the properties or behavior of the system under study ([13, sec 2]). It is modeled probabilistically by a random vector $a \in \mathbb{L}^0(\Omega_a, \mathbb{R}^{d_a})$.

The aim of this modeling is to have distinct representations of these two sources of uncertainty to see what confidence can be put in the value of a margin, instead of aggregating everything.

In these works, the margin is defined as the difference between a requirement R placed on the performance of a system and the predicted performance P [13, sec 6.1]. In general, R and P are indicators of the same dimension d calculated by aggregating the aleatory uncertainty but still carrying the epistemic uncertainty i.e they are random variables of the space $\mathbb{L}^0(\Omega_e, \mathbb{R}^d)$. For instance, P can be a conditional expectation with respect to e of a function of a and e of the form $\mathbf{E}[f(a, e)|e]$.

The use of epistemic uncertainty is motivated by the fact that the requirement R is not always precisely specified and the underlying phenomenon modeled by P is not fully characterized. In this context, a margin has the form M(R, P) with M(r, p) a function taking deterministic inputs which is nonnegative if and only if the performance p meets the requirement r (for example $p \leq r$ when these two quantities are scalar). M(R, P) is then also a random variable of $\mathbb{L}^0(\Omega_e, \mathbb{R})$.

Conditionally to the epistemic uncertainty, this margin definition is a generalization of the margin on the risk level as it is a difference between two statistical indicators. For instance in [15], some margins are formulated as $M = E_{req} - \mathbf{E} \left[f(a, e) | e \right]$ with $R = E_{req}$ being deterministic. One of the interest in keeping the epistemic randomness is to model the confidence one has on the margin value via a random variable. It can then be represented via graphical tools, such as a cumulative distribution function plot, as shown in [15, Figure 2.6].

To describe it in our framework, we use a random model of margin on a probability space $(\Omega_a, \mathscr{F}_a, \mathbf{P}_a)$, as described in Definition 2.9. We furthermore impose that the model of margin M and each point $U \in \mathscr{C}$ are functions of a vector e, so that the margin $\mathbf{m}(U(e); \mathbf{M}(e))$ is itself a function of e. Then, we model the epistemic uncertainty by choosing a law of probability for e, on another probability space $(\Omega_e, \mathscr{F}_e, \mathbf{P}_e)$. Intuitively, U(e) models the uncertainties on the studied state of the studied phenomenon whereas $\mathbf{M}(e)$ models the uncertainties on the model/equations via $\mathscr{C}(e)$ and on the requirements via $\mathscr{A}(e)$. The margin is a random variable:

$$\mathsf{m}\left(U(e);\mathsf{M}(e)\right) \in \mathbb{L}^{0}\left(\Omega_{e},\mathbb{R}\right).$$
(33)

An interesting property of such an approach lies in the fact that m(U(e); M(e)) is now a *real* random variable and thus is easier to study. It enables us to interpret the confidence on a value of margin as well as plotting its cumulative distribution function.

4.6.3 Margin determination and allocation in complex systems

Classic margin problems in the field of complex multidisciplinary systems design are called *margin determination* or *margin allocation*. Some instances of these problems can be seen as determining the optimal requirements on the system that should be asked at a given phase of design, while keeping a reserve to cover some uncertainties in other phases, in the spirit of the margin defined in [23]. Margins are used in this context as a way of fixing the reserve values, as well as tools to keep track of the evolution of the system and trade a parameter performance over another. This section explores links to works on margins in this context to our model of margin.

A margin definition of the type "the lower/the greater the better" In this context, the existing works that we found focused on margins on parameters for which the risk was monotonic [12, Section 2 B], which is particularly suitable to study small systems. More precisely, if we denote by u a real variable of interest, it is possible to define a margin only if it falls in the following category:

- The lower the better, a smaller value of u leads to a lower risk, all other parameters being equal.
- The greater the better, a greater value of *u* leads to a lower risk, all other parameters being equal.
- The closest to a value the better a value of u closer to a known constant u_0 leads to a lower risk, all other parameters being equal.

Method for Determining Margins in Conceptual Design A framework to determine margins is described in the work of D.P Thunissen in [29, 30]. Here, determining margins means determining the value of the demanded margin for a given model of margin, i.e calibrating the model of margin as it will be described in Section 5.1. If we focus on margins, the proposed method can be seen as:

- 1. Identifying the important tradable performance parameters y_1, \ldots, y_n [30, Sec 4.1]. It typically involves measures of the system performances mass of a component, power required...-, costs, and schedule duration.
- 2. Modeling these performance parameters with random variables Y_1, \ldots, Y_n , propagating the uncertainty sources that have been identified. This model can be reevaluated as more information is available.
- 3. Choosing the requirements on the performance parameters so as to cover a proportion r of the best cases. r can be chosen equal to 0.90, 0.95 or 0.99 depending on the criticality of these parameters. These requirements, expressed as the quantiles $q_{r_1}(Y_1), \ldots, q_{r_n}(Y_n)$, are computed thanks to the probabilistic model.

If we denote by q(Y) the limit imposed as a quantile on the parameter y in the step 3 of the process, the margin they propose on y is:

$$m_{
m demanded} = q(Y) - y_{
m det},$$

 $m_{
m demanded,norm} = rac{q(Y) - y_{
m det}}{y_{
m det}},$

with y_{det} being designated as the "deterministic value" of y and seems to be either the best current estimate of y at a given moment of the project or the chosen design value *before* imposing a margin.

First, let us remark that this definition of margin prescribes "how much extra quantity should be taken". It is implicitly implied that this extra quantity is taken from the reference of y_{det} in the direction of an increase of y - we are in a *the lower the better* case:

$$\mathsf{m}\left(y;\mathsf{M}\right) = y - y_{\mathrm{det}},$$

or in the normalized case:

$$\mathsf{m}\left(y;\mathsf{M}_{\text{normalized}}\right) = \frac{y - y_{\text{det}}}{y_{\text{det}}}$$

Their approach can be summarized as imposing a demanded margin on the design with a value of

$$\mathbf{m}(y; \mathbf{M}) \ge m_{\text{demanded}}.$$
 (34)

Modeling y by a random variable Y and using it to choose the margin value is a margin calibration, as described in Section 5.1. Even if this method can be equivalently reformulated as a condition on a quantile of the form y > q(Y), the interest of such an approach is:

- Once q(Y) has been computed, the margin is deterministic, interpretable and can be compared to previous margin values. They are used to perform trade-offs between parameters.
- The random variable Y can change with architecture changes or the information available over the time. It is then interesting to update the required margin value and to compare it with previous values.

A framework to see the cost and allocate margins in complex system design More recent works have included this method in a wider framework to design complex systems [12, Section III]. One of their main contributions is the definition of a margin space which considers the demanded margins on design parameters as additional parameters in the design space. Choosing a margin $m_{demanded,y}$ on y virtually translates the coordinate y of the design point of $m_{demanded,y}$ in the direction (either negative or positive) of an increase in risk. This new point is used as a reference to check if the constraints are still satisfied. The direction is determined by the nature of the parameter "the lower the better" or "the greater the better". For instance if y is of type "the lower the better", then a margin would translate the design point to $y + m_{demanded,y}$ and the requirement would be fulfilled for this value.

By doing so, it is possible to visualize the effect that imposing a margin value has on the design space and on the performances; it is called *Margin vs Performances* [12, Section III B 3.]. Indeed, imposing greater margins reduces the possible design and has an impact on the performance. In a certain sense, this shows the *cost* of a margin.

Another interesting application is the possibility to represent the random model Y of the parameter y (designed in Step 2 of the previous process) in the design space. It is possible to choose the margin so that satisfying the deterministic constraints would lead to a high enough probability of constraint satisfaction. This is called in these works *Tradeoff between margins and Probability of Constraint Satisfaction* [12, Section III B 2. - Figure 9] and this is what we call the *Margin calibration* in Section 5.1.

Let us remark that the underlying models of margins used here are of the type always acceptable box (defined in Section 2.2.2). Indeed, the parameters are of the type "the lower the better" or "the greater the better". When a margin $m_{\text{demanded},y}$ is imposed on y, it is assumed that the constraints will be satisfied for all values in the interval with margin $[y - m_{\text{demanded},y}; y + m_{\text{demanded},y}]$, as long as the other parameters also belong to their respective interval with margin. Then the normalized margins can be seen as imposed safety coefficients, studied in Section 4.3. Using a wider class of models of margin proposed in this paper could help to overcome at least two problems:

- How to define margins and impose margins on a parameter that is not of the type "the lower the better", "the greater the better" or "the closer to a known value the better"?
- How to take a margin on multiple parameters that is not of the form of an always acceptable box, i.e that is not only operating ranges for each variables?

The margins could then be defined to meet the designer expectations more precisely as well as continuing to be an interpretable real number with a physical unit (or a percentage).

5 Complementary remarks

Before concluding on this work, we emphasize some aspects of the margin that the model of margin does not cover directly but that are often a part of the process in the margin usage. They can be seen as some of the possible following steps, after the definition of the model of margin.

5.1 Margin calibration

As stated in the introduction, a margin is taken to prevent some unwanted effects that are uncertain. How to assess these uncertainties and their consequences is not in the scope of this paper and must be a part of the problem modeling. However, it is possible to formulate a problem of margin calibration, once the uncertainties and consequences have been identified. Informally, a problem of margin calibration is:

Definition 5.1 (Margin calibration - Informal). Find the value $m_{\lim} \ge 0$ and a margin model M such that if $m(U; M) \ge m_{\lim}$ then at the design point U all the uncertainties/consequences that one wants to take into account are covered.

Two strategies are used in practice for this purpose:

• Develop another model on U, denoted by U' and an associated risk measure and threshold μ, μ_{req} , which takes into account the uncertainties/consequences that are not modeled in U. It can be a modeling uncertainty, an epistemic uncertainty, an estimation by the experts of the "unknown unknowns" for instance. In that case, the margin calibration is:

Definition 5.2 (Margin calibration - external model). Find $m_{\text{lim}} \ge 0$ such that if $m(U; M) \ge m_{\text{lim}}$ then $\mu(U') < \mu_{\text{req}}$.

This approach is used to choose the values of the Safety Coefficients (Section 4.3) and is called semiprobabilistic calibration in [20, Section 10.1.4]. It is mentioned in the Phase/Gain Margin method [21, Section 8.4.2]. It is at the heart of the margin allocation/determination framework studied in Section 4.6.3.

• Model in \mathcal{A} all the uncertainties/consequences that must be covered. In other terms, all the uncertainties are in the constraints. The margin calibration is formulated in this case:

Definition 5.3 (Margin calibration - Acceptance set). Find the acceptance set \mathscr{A} for the margin model M such that $m(U; M) \ge 0$ implies that all uncertainties/consequences are covered.

The Capability process index (Section 4.1) is an example of such an approach, it is calibrated so that a C_p greater than 1 (and thus a margin greater than 0) leads to a probability of unwanted event of 0.001. This approach also arises when the designer follows some regulatory requirements: in that case, the acceptance set (for instance $\mathscr{A} = \{x > x_{req}\}$) might include all the uncertainty that is necessary to take into account.

First, let us remark that these two approaches are somewhat mathematically equivalent: imposing a margin $m(U; M) > m_{\lim}$ can be equivalently modeled with a model of margin with acceptance set $\mathscr{A} = \{U | m(U; M) > m_{\lim}\}$. The difference lies in the meaning that is given to the margin values and the acceptance set in engineering practices.

This explains for instance why, in some practices, a margin of value 0 is perfectly admissible - the uncertainties/consequences are all in \mathcal{A} - but in other practices it is unconceivable - a null margin would mean that the system is very close to a failure point.

We can reasonably imagine that these two approaches are often both used at the same time - although sometimes implicitly - in the process of a margin definition. In the context of a design analysis, having this in mind can help mapping where the actual margins are and what uncertainties/consequences they really cover. It also gives some insights on how to define and describe the link between the margin and the uncertainties/consequences.

5.2 Margin allocation

The problem of margin allocation arises in a multidisciplinary design context, when one needs to specify some acceptable values on a variable of interest u, for other disciplines to meet this criterion. Let us take for instance the case of a deterministic additive model in which the state space is $\mathbf{E} = \{u = (y, x_1, x_2) \in \mathbb{R}^3\}$ and $\mathscr{C} = \{(y, x_1, x_2) \in \mathbb{E} | y = x_1 + x_2\}$. Let us assume furthermore that y must belong to $[y_{\text{ref}} - \Delta y, y_{\text{ref}} + \Delta y]$. The acceptance set is then $\mathscr{A} = \{(y, x_1, x_2) \in \mathbb{E} | y \in [y_{\text{ref}} - \Delta y, y_{\text{ref}} + \Delta y]\}$.

A typical margin allocation problem would be to specify some acceptable intervals for x_1 and x_2 so that the condition on y would always be satisfied. These intervals would then be given to other disciplines that would be in charge to find some values of x_1 and x_2 that meet these requirements. All intervals of the form $x_1 \in [x_{1,\text{ref}} - \Delta x_1; x_{1,\text{ref}} + \Delta x_1]$ and $x_2 \in [x_{2,\text{ref}} - \Delta x_2; x_{2,\text{ref}} + \Delta x_2]$ with $|x_{2,\text{ref}} + x_{1,\text{ref}} - y_{\text{ref}}| \leq \Delta y - \Delta x_2 - \Delta x_1$ would be acceptable. The problem comes from the fact that small values of Δx_1 and Δx_2 could be expensive because the requirements are too strict. If it is possible to model all the costs by a real function $C(\Delta x_1, \Delta x_2)$, we can formulate the allocation problem as:

$$\min C(\Delta x_1, \Delta x_2)$$

s.t $|x_{2,ref} + x_{1,ref} - y_{ref}| \le \Delta y - \Delta x_2 - \Delta x_1$ (35)

In engineering practices, the interval $[x_{1,ref} - \Delta x_1; x_{1,ref} + \Delta x_1]$ is sometimes called a margin, passed to another discipline in charge of this quantity. This interval is indeed a demanded margin for the following margin model M₁: E, \mathscr{C} , \mathscr{A} are as defined above. The probed coordinates are the evolution in x_1 , for $u = (y, x_1, x_2)$:

$$\mathscr{G}_{u|\mathscr{C}} = \left\{ \left(y', x_1', x_2'\right) \in \mathscr{C} | x_2' = x_2, x_1' = x_1 + k, k \in \mathbb{R} \right\}$$

The coordinate of interest $\mathbf{S} = \mathbb{R}$ is the change with respect to x_1 and the coordinate functions are

$$\phi_u(y', x_1', x_2') = |x_1' - x_1|$$

Our model of margin M_1 is now complete. For $u = (x_{1,\text{nom}} + x_{2,\text{nom}}, x_{1,\text{nom}}, x_{2,\text{nom}})$, if $\mathbf{m}(u; M_1) \ge m_{x_1}$, the acceptability property gives that for each $|\Delta x| \le m_{x_1}$, the point $(x_{1,\text{nom}} + \Delta x + x_{2,\text{nom}}, x_{1,\text{nom}} + \Delta x, x_{2,\text{nom}})$ is in the acceptance set \mathscr{A} . In other terms, it is sufficient to impose:

$$x_1 \in [x_{1,\text{nom}} - m_{x_1}; x_{1,\text{nom}} + m_{x_1}]$$

as a requirement for the design to be acceptable.

The demanded margin m_{x_1} represents a requirement that the system withstand a variation of x_1 at the point u. This is the point of view "extra quantity included" developed in this article. The dual point of view is that, when passed as a requirement to another discipline, m_{x_1} can be interpreted as a room for maneuvre. It is a measure of the freedom for the discipline in charge of fixing x_1 . This explains why the word "margin" is used in both acceptions, depending on the context.

In order to choose the demanded margin m_{x_2} for x_2 , (at least) two strategies can be exhibited. One to fix $m_{x_1} > 0$ first and then choose m_{x_2} with what remains. In that case, the margin model M₂ would be defined in a similar way to M₁ replacing x_1 by x_2 , except that the acceptance set would be all the points having already at least m_{x_1} of margin:

$$\mathscr{A}_2 = \left\{ u \in \mathbf{E} | \mathsf{m}\left(u; \mathsf{M}_1\right) \ge m_{x_1} \right\}.$$

Another possibility would be to choose a model of margin of type always acceptable box on x_1 and x_2 , defined in Section 2.2.2.

As far as the first method is concerned, if we denote the cost associated to the margins by $C(m_{x_1}, m_{x_2})$, we can reformulate the allocation problem at a given point u as:

min
$$C(m_{x_1}, m_{x_2})$$

s.t m $(u; M_1) \ge m_{x_1}$ and m $(u; M_2) \ge m_{x_2}$

In practice, this cost function might no be well defined, but the models of margin M_1 and M_2 can often be exhibited. The choice of the values m_{x_1} and m_{x_2} would be done by expert judgment.

This formulation could possibly be used in more complicated models, e.g with more variables, nonlinear equations or models of margin with random variables. Other possibilities might also include a change in the design point u to have better allocation possibilities.

6 Conclusion

In this paper, we have proposed two mathematical definitions to formalize the concept of margin, understood as an amount of something included so as to be sure of success or safety. The first one is the model of margin, given in Definition 2.4, that describes all the necessary components to compute a margin. It consists in the description of the problem, the modeling of the phenomenon and the accepted states in $\mathbb{E}, \mathcal{C}, \mathcal{A}$. The probing sets $(\mathcal{G}_u)_{u \in \mathcal{C}}$ define the possible evolutions of the phenomenon taken into account in the computation of the margin. The coordinates of interest $\mathbf{S}, d_{\mathbf{S}}$ define on which coordinates to compute a margin. The second one, given in Definition 2.5, tells how to compute the margin at a given point u, for a given model of margin. The margin and model of margin can also be used in a probabilistic modeling context, as shown in Definitions 2.9 and 2.10. The acceptance set \mathcal{A} can be defined thanks to a risk measure, this point is discussed in Section 3.1.

Then, we have exhibited the possible uses of such a formalism. In Section 4, we have shown that some classical indicators of the literature that can be intuitively interpreted as a margin can be rewritten as models of margin. This enables to unify their definitions and to have a better understanding of their roles looking at their models of margin. In Section 5, we have put in light the fact that this formalism can be used to give an abstract formulation of two classes of problems in engineering under uncertainty, independently from the field considered. The problem of margin calibration consists in choosing the minimal value of the margin so as to cover uncertainties and their related consequences. The problem of margin allocation consists in giving some specifications on other interacting engineering disciplines, using the margin formalism.

References

- P. ARTZNER, F. DELBAEN, J.-M. EBER, AND D. HEATH, Coherent measures of risk, Mathematical finance, 9 (1999), pp. 203–228.
- [2] J. R. BENJAMIN AND C. CORNELL, Probability, statistics and decision for civil engineers, 1970.
- [3] E. C. M. DE WECK OLIVIER, C. P. JOHN, ET AL., A classification of uncertainty for early product and system design, Guidelines for a Decision Support Method Adapted to NPD Processes, (2007), pp. 159–160.

- [4] H. DEZFULI, A. BENJAMIN, C. EVERETT, G. MAGGIO, M. STAMATELATOS, S. GUARRO, P. RUT-LEDGE, J. SHERRARED, C. SMITH, AND R. WILLIAMS, *Nasa risk management handbook*, tech. rep., 11 2011.
- [5] DOE-NE, Doe-ne light water reactor sustainability program and epri long term operations program joint research and development plan, Idaho National Laboratory, (2015).
- [6] C. ECKERT AND O. ISAKSSON, Safety margins and design margins: A differentiation between interconnected concepts, Procedia CIRP, 60 (2017), pp. 267 – 272. Complex Systems Engineering and Development Proceedings of the 27th CIRP Design Conference Cranfield University, UK 10th – 12th May 2017.
- [7] C. ECKERT, O. ISAKSSON, AND C. EARL, Design margins: a hidden issue in industry, Design Science, 5 (2019), p. e9.
- [8] Evaluation of measurement data Guide to the expression of uncertainty in measurement., Tech. Rep. ISO/IEC Guide 98-3, JCGM, 1995.
- [9] H. FÖLLMER AND A. SCHIED, Stochastic finance: an introduction in discrete time, Walter de Gruyter, 2011.
- [10] G. F. FRANKLIN, J. D. POWELL, AND A. EMAMI-NAEINI, Feedback control of dynamic systems, Addison-Wesley Reading, MA, sixth ed., 2010.
- [11] J.-Y. GOTOH, A. TAKEDA, AND R. YAMAMOTO, Interaction between financial risk measures and machine learning methods, Computational Management Science, 11 (2014), pp. 365–402.
- [12] M. D. GUENOV, X. CHEN, A. MOLINA-CRISTÓBAL, A. RIAZ, A. S. J. VAN HEERDEN, AND M. PADULO, Margin allocation and tradeoff in complex systems design and optimization, AIAA Journal, 56 (2019), pp. 2887–2902.
- [13] J. C. HELTON, Quantification of margins and uncertainties: Conceptual and computational basis, Reliability Engineering & System Safety, 96 (2011), pp. 976 – 1013. Quantification of Margins and Uncertainties.
- [14] J. C. HELTON AND J. D. JOHNSON, Quantification of margins and uncertainties: Alternative representations of epistemic uncertainty, Reliability Engineering & System Safety, 96 (2011), pp. 1034–1052.
- [15] J. C. HELTON, J. D. JOHNSON, AND C. J. SALLABERRY, Quantification of margins and uncertainties: example analyses from reactor safety and radioactive waste disposal involving the separation of aleatory and epistemic uncertainty, Reliability Engineering & System Safety, 96 (2011), pp. 1014–1033.
- [16] S. HESS, J. GAERTNER, J. GABOR, L. SHANLEY, L. ENOS-SYLLA, S. PRASAD, AND S. HOLLINGSWORTH, Framework for risk-informed safety margin characterization, EPRI Report, 1019206 (2009).
- [17] M. HREHOR, M. GAVRILAS, J. BELAC, R. SAIRANEN, G. BRUNA, M. REOCREUX, F. TOUBOUL, B. KRZYKACZ-HAUSMANN, J. S. PARK, A. PROSEK, ET AL., Task group on safety margins action plan (smap). safety margins action plan-final report, tech. rep., Organisation for Economic Co-Operation and Development, 2007.
- [18] Statistical methods in process management capability and performance part 4: Process capability estimates and performance measures, tech. rep., International Organization for Standardization, 2016.
- [19] Statistical methods in process management capability and performance –, tech. rep., International Organization for Standardization, 2016.
- [20] M. LEMAIRE, Structural reliability, John Wiley & Sons, 2013.
- [21] W. S. LEVINE, *The Control Systems Handbook: Control System Fundamentals*, CRC press, second edition ed., 2010.
- [22] F. MANGEANT, Survey of industrial practices around UQ: state of the art & challenges, in Proceedings of UQOP : Uncertainty Quantification and OPtimization, 2019.
- [23] H. MCMANUS AND D. HASTINGS, A framework for understanding uncertainty and its mitigation and exploitation in complex systems, in INCOSE International Symposium, vol. 15, Wiley Online Library, 2005, pp. 484–503.
- [24] D. C. MONTGOMERY, Introduction to statistical quality control, John Wiley & Sons, 2007.
- [25] R. T. ROCKAFELLAR, Coherent Approaches to Risk in Optimization Under Uncertainty, ch. Chapter 3, pp. 38–61.
- [26] R. T. ROCKAFELLAR AND J. O. ROYSET, Engineering decisions under risk averseness, ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, 1 (2015), p. 04015003.

- [27] R. R. SHERRY, J. R. GABOR, AND S. M. HESS, Pilot application of risk informed safety margin characterization to a total loss of feedwater event, Reliability Engineering & System Safety, 117 (2013), pp. 65–72.
- [28] C. M. C. J. SLUTSKY AND D. ANTIS, Design for Six Sigma in Technology and Product Development., Prentice Hall Professional, 2003.
- [29] D. P. THUNNISEN, Method for determining margins in conceptual design, Journal of spacecraft and rockets, 41 (2004), pp. 85–92.
- [30] D. P. THUNNISSEN, Propagating and mitigating uncertainty in the design of complex multidisciplinary systems, PhD thesis, California Institute of Technology, 2005.
- [31] J. VON NEUMANN, O. MORGENSTERN, ET AL., Theory of games and economic behavior, (1953).
- [32] R. YOUNGBLOOD, Risk-informed safety margin characterization (rismc): Integrated treatment of aleatory and epistemic uncertainty in safety analysis, tech. rep., Idaho National Laboratory (INL), 2010.