Fast Belief Estimation in Evidence Network Models

Massimiliano Vasile
Professor, Aerospace Centre of Excellence, Mechanical and Aerospace Engineering
University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, UK
Email: massimiliano.vasile@strath.ac.uk

Gianluca Filippi
Postgraduate Researcher, Aerospace Centre of Excellence, Mechanical and Aerospace Engineering
University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, UK
Email: gian.filippi.91@gmail.com

Carlos Ortega Absil
PhD candidate, Aerospace Centre of Excellence, Mechanical and Aerospace Engineering
University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, UK
Email: carlos.ortega@strath.ac.uk

Annalisa Riccardi
Lecturer, Aerospace Centre of Excellence, Mechanical and Aerospace Engineering
University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, UK
Email: annalisa.riccardi@strath.ac.uk

Summary

This paper introduces a novel approach to model complex engineering systems in the form of a network of interconnected nodes. Each node is associated to a value and an evidence measure associated to that value. The Belief and Plausibility associated to the total value of the network is then estimated with a fast decomposition technique that allows for several order of magnitude reduction in computational time under some assumptions on the properties of the network. The modeling approach and associated Belief estimation technique are proposed for the optimisation of complex engineering systems under epistemic uncertainty. The methodology is applied to the preliminary design of a small satellite where some quantities are affected by an epistemic uncertainty. In addition, the paper describes a surrogate method that provides a faster evaluation of the belief curve.

Keywords: *Evidence, Belief, Decomposition, Optimization.*

1 Introduction

The quantification of uncertainty in complex engineering systems is computationally challenging. Part of the challenge comes from the computational cost of propagating uncertainty through the system model. This cost is typically exponential with problem dimension and can become quickly untreatable even for systems of moderate size. Uncertainty can be modeled as a random process with an associated probability distribution. In a number of cases, however, the knowledge on the probability distribution is missing or is uncertainty in itself. In this later case the uncertainty is epistemic as there is a fundamental lack of knowledge or information.

The paper proposes to model a complex engineering system as a particular network of interconnected components. Each component is exchanging information with other components through a set of exchange variables. The quantity associated to each component and the exchange variables are assumed to be affected by uncertainty, where this uncertainty can be epistemic in nature. In particular, we propose the use of Dampster-Shafer theory of evidence to model epistemic uncertainty.[Shafer(1976a)] For this reason the network model representing the engineering system is here called Evidence Network Model (ENM). Evidence theory, that belongs to a class of mathematical theories known as Imprecise Probabilities, can be seen as an extension to standard probability theory and is devised to properly handle the imprecision that comes with epistemic uncertainty.

In this paper we consider the case in which an ENM is characterised by a global quantity of interest F:

$$F : D \times U \subseteq \mathbb{R}^{m+n} \to \mathbb{R} \tag{1}$$

where F depends on some uncertain parameters $\mathbf{u} \in U \subset \mathbb{R}^m$ and design parameters $\mathbf{d} \in D \subset \mathbb{R}^n$. The set D is the available design space and U the uncertain space. One or more intervals are defined for each uncertain variable u_i and for each one of those intervals a level of confidence on the values of u (bpa, basic probability assignment) is assigned. All the Cartesian products of the uncertain intervals with a non-zero basic probability assignment are called Focal Elements (FE). Assuming strong independence among uncertain variables the bpa associated to each FE is the product of the bpas associated to each interval. For a given design $\mathbf{d}$, we can then calculate two cumulative quantities, Belief (Bel) and Plausibility (Pl),[Shafer(1976b)] that provide the lower and upper belief in the occurrence of a particular event or the truth of a proposition on the value of the network.

Once the ENM is defined, a decomposition technique is used to quickly quantify the uncertainty associated to F. The decomposition approach, proposed in this paper, starts from some suitable assumptions on the nature of the exchange variables and the dependency of the quantity of interest F on these variables. It then proceeds by decomposing the ENM and incrementally reconstructing Belief and Plausibility at a computational cost that is orders of magnitude lower than a full computation of Belief and Plausibility.[Alicino and Vasile(2014a)]

The paper presents some illustrative examples: three synthetic function and one real problem, the design of a small satellite in which three subsystems are affected by epistemic uncertainty.

2 Evidence Network Models

A generic complex system can be represented as a network, where each node represents one of its subsystem and links represents sharing of information between subsystems. We can then define a function F as

$$F(\mathbf{d},\mathbf{u}) = \sum_{i=1}^{N} g_i(\mathbf{d},\mathbf{u}_i,\mathbf{h}_i(\mathbf{d},\mathbf{u}_i,\mathbf{u}_{ij}))$$

where N is the number of subsystems involved, $\mathbf{h}_i(\mathbf{d},\mathbf{u}_i,\mathbf{u}_{ij})$ is the vector of scalar functions $h_{ij}(\mathbf{d},\mathbf{u}_i,\mathbf{u}_{ij})$ where $j \in J_i$ and J_i is the set of indexes of nodes connected to the i-th node; $\mathbf{u}_i$ are the uncertain variables of subsystem i not shared with any other subsystems and $\mathbf{u}_{ij}$ are the uncertain variables shared among subsystems i and j. Please note that accordingly to our notation $\mathbf{u}_{ij} = \mathbf{u}_{ji}$ and the functions $g_i(\cdot,\cdot,\cdot)$ represent quantities computing by the governing equations of the different subsystems. In a fully connected network as in Figure 1 the function F is:

$$\begin{aligned} F(\mathbf{d},\mathbf{u}) \quad = \quad & g_1(\mathbf{d},\mathbf{u_1},h_{12}(\mathbf{d},\mathbf{u_1},\mathbf{u_{12}}),h_{13}(\mathbf{d},\mathbf{u_1},\mathbf{u_{13}})))+ \\ & g_2(\mathbf{d},\mathbf{u_2},h_{21}(\mathbf{d},\mathbf{u_2},\mathbf{u_{12}}),h_{23}(\mathbf{d},\mathbf{u_2},\mathbf{u_{23}})))+ \\ & g_3(\mathbf{d},\mathbf{u_3},h_{31}(\mathbf{d},\mathbf{u_3},\mathbf{u_{13}}),h_{32}(\mathbf{d},\mathbf{u_3},\mathbf{u_{23}})). \end{aligned} \tag{2}$$

We then call $\mathbf{u}_i$, *uncoupled* variables because they influence only one subsystem and $\mathbf{u}_{ij}$ *coupled* variables because they influence two subsystems. Hence for the example in Figure 1 the uncertain vector can be ordered as

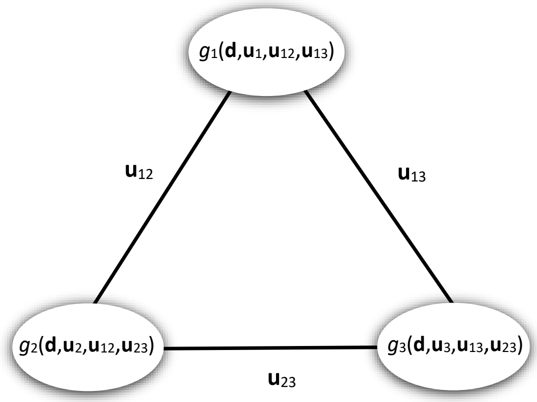

Figure 1: Evidence Network Model of a generic system F composed of three sub-systems with coupled variables u_{12}, u_{13} and u_{23}.

$$\mathbf{u} = [\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3,\mathbf{u}_{12},\mathbf{u}_{13},\mathbf{u}_{23}]^T.$$

In the following we will study only the case in which the functions $g_i(\cdot,\cdot,\cdot)$ are always positive semidefinite and are monotonic with respect to each function h_{ik}.

Given a design, or decision, value $\tilde{\mathbf{d}} \in$ D we will call *worst case scenario* the vector $\underline{\mathbf{u}}$ that corresponds to the maximum of F over the space U:

$$\underline{\mathbf{u}} = \underset{\mathbf{u}\in U}{\operatorname{argmax}} F(\widetilde{\mathbf{d}},\mathbf{u}) \tag{3}$$

likewise we can call *best case scenario* the quantity:

$$\bar{\mathbf{u}} = \underset{\mathbf{u}\in U}{\operatorname{argmin}} F(\widetilde{\mathbf{d}},\mathbf{u}) \tag{4}$$

We can now define an event in the space U, or a proposition on the value of F, as the set A such that:

$$A = \{\mathbf{u} \in U \,|\, F(\mathbf{d},\mathbf{u}) \leq \nu\} \tag{5}$$

From this definition it is clear that for every design $\mathbf{d} \in D$ the worst case scenario corresponds to $A = U$, because $\nu = \max_{\mathbf{u}\in U} F(\mathbf{d},\mathbf{u})$, and analogously the best case scenario has zero measure. Each uncoupled uncertain vector $\mathbf{u}_i$ is defined over a set of boxes named $\Theta_i = \cup_k \theta_{k,i}$ and each coupled uncertain vector $\mathbf{u}_{ij}$ is defined over the set of boxes $\Theta_{ij} = \cup_k \theta_{k,ij}$. We define the set

$$\Theta = \bigcup_i \theta_i = (\times_{i=1}^{m_u} \Theta_i) \times (\times_{i,j=1}^{m_c} \Theta_{ij})$$

where m_u is the number of uncoupled uncertain vectors (equal to the number of subsystems) and m_c is the number of coupled uncertain vectors. and the hyperpower set

$$D^\Theta = (\Theta, \cup, \cap) \tag{6}$$

as the set composed of the elements of Θ, their union and intersection. In the following the space $U := D^\Theta$. We can then define quantities associated to the belief in the occurrence of the event A:

$$Bel(A) = \sum_{\theta\subset A, \theta\in U} bpa(\theta) \tag{7}$$

$$Pl(A) = \sum_{\theta\cap A\neq 0, \theta\in U} bpa(\theta) \tag{8}$$

where $bpa(\theta)$ is the basic probability assignment associated to θ, an element of the power set. It is important to note that if the h_{ij} functions were known with certainty the nodes composing the network would be decoupled and statistically independent. We also note that in order to identify if a θ is fully included in A we need to find the maximum of F with respect to $\mathbf{u} \in \theta$. Likewise an intersection with A requires computing the minimum of F with respect to $\mathbf{u} \in \theta$. Given that the subsets θ, their unions and intersections come from a cross product, it is clear that the number of maximisation and minimisation increases exponentially with the number of dimensions. The computation of the Belief in the occurrence of A is, therefore, an exponentially complex operation. In the following section a technique is proposed to compute an approximation to (7) by exploiting some of the properties of the ENM listed above. In particular we will exploit the following three properties:

1. The contribution of the coupled varaible $\mathbf{u}_{ij}$ to the value F manifests through the scalar functions h_{ij} and h_{ji}.
2. All g_i functions are positive semidefined.
3. All g_i functions are monotonically increasing with respect to h_{ij} for every j.

3 Decomposition Algorithm

In order to reduce the computational complexity of the calculation of $Bel(A)$ we propose a decomposition technique that exploits the three properties defined in the previous section. The decomposition algorithm aims at decoupling the sub-systems over the uncertain variables in order to optimise only over a small sub-set of the Focal Elements (Algorithm 1); this procedure requires the following steps:

1. Solution of the optimal worst case scenario problem:
$$\min_{\mathbf{d}\in D} \max_{\mathbf{u}\in U} F(\mathbf{d},\mathbf{u}) \tag{9}$$
2. Maximisation over the coupled variables and computation of $Bel(A)$.
3. Maximisation over the uncoupled variables.
4. Reconstruction of the approximation $\widetilde{Bel}(A)$.

Point 1 will not be discussed in this paper. An algorithm can be found in,[Vasile(2014)Alicino and Vasile(2014b)] and more recently in.[Ortega and Vasile(2017)] The result of the solution of problem (9) are the values $\tilde{\mathbf{d}}$ and $\underline{\mathbf{u}}$, thus, in the following the assumption is that $\tilde{\mathbf{d}}$ is already available.

3.1 Maximisation over the coupled variables and evaluation of the partial Belief curves

For each coupled vector $\mathbf{u}_{ij}$ a maximisation is run over each Focal Element $\theta_{k,ij} \subseteq \Theta_{ij} \subseteq U$, given $\widetilde{\mathbf{d}}$ and keeping fixed all the other components of $\mathbf{u}$. Taking again the example in Figure 1 we have:

$$\begin{aligned}
\hat{\mathbf{u}}_{k,12} &= \operatorname*{argmax}_{\mathbf{u}_{12}\in\theta_{k,12}} F(\widetilde{\mathbf{d}},\underline{\mathbf{u}}_1,\underline{\mathbf{u}}_2,\underline{\mathbf{u}}_3,\mathbf{u}_{12},\underline{\mathbf{u}}_{13},\underline{\mathbf{u}}_{23}), \forall\theta_{k,12}\subset\Theta_{12}\\
\hat{\mathbf{u}}_{k,13} &= \operatorname*{argmax}_{\mathbf{u}_{13}\in\theta_{k,13}} F(\widetilde{\mathbf{d}},\underline{\mathbf{u}}_1,\underline{\mathbf{u}}_2,\underline{\mathbf{u}}_3,\underline{\mathbf{u}}_{12},\mathbf{u}_{13},\underline{\mathbf{u}}_{23}), \forall\theta_{k,13}\subset\Theta_{13}\\
\hat{\mathbf{u}}_{k,23} &= \operatorname*{argmax}_{\mathbf{u}_{23}\in\theta_{k,23}} F(\widetilde{\mathbf{d}},\underline{\mathbf{u}}_1,\underline{\mathbf{u}}_2,\underline{\mathbf{u}}_3,\underline{\mathbf{u}}_{12},\underline{\mathbf{u}}_{13},\mathbf{u}_{23}), \forall\theta_{k,23}\subset\Theta_{23}
\end{aligned} \tag{10}$$

For easiness in the notation we will indicate with

$$F(\mathbf{u}_{ij}) := F(\tilde{\mathbf{d}},\underline{\mathbf{u}}_1,...,\mathbf{u}_{ij},...,\underline{\mathbf{u}}_{i+1j},...).$$

We can then compute the partial belief associated only to the coupled variables with index ij:

$$Bel(F(\mathbf{u}_{ij}) < \nu) = \sum_{\theta_{k,ij}|\max_{\mathbf{u}_{ij}\in\theta_{k,ij}} F(\mathbf{u}_{ij})\leq\nu} bpa(\theta_{k,ij}) \tag{11}$$

The calculation of the partial belief can be found in Algorithm 1, line 6. Once the partial belief curve, for each coupled vector, is available, one can sample these curves, by taking a succession of $\{\nu_1,...,\nu_{N_S} = \nu\}$ values, and find the corresponding values of the coupled vectors $\widehat{\mathbf{u}}_{k,ij}^q$. These values will be used in the next step to decouple the functions g_i (g_j) and compute the maxima of each g_i (g_j) with respect to the uncoupled variables $\mathbf{u}_i$ ($\mathbf{u}_j$).

3.2 Optimization over the uncoupled vectors

For each level q, given a fix value of the coupling functions, one can study each g_i independently of the others. The idea is to run an optimisation for each function g_i over only the uncoupled vector $\mathbf{u}_i$. With the example in Figure 1 in mind, having

$$\hat{h}^q_{ij}(\mathbf{u}_i) := h_{ij}(\tilde{\mathbf{d}}, \mathbf{u}_i, \hat{\mathbf{u}}^q_{ij})$$

where $\hat{\mathbf{u}}^q_{ij} := \hat{\mathbf{u}}^q_{k^*,ij} : k^* = \mathrm{argmax}_k F(\hat{\mathbf{u}}^q_{k,ij})$, is one of the maxima of the maxima attained by the coupled variable $\mathbf{u}_{ij}$. For every Focal Element $\theta_{k,i} \in \Theta_i$ we have:

$$\begin{aligned}
\hat{\mathbf{u}}^q_{k,1} &= \underset{\mathbf{u}_1 \in \theta_{k,1}}{\mathrm{argmax}}\, g_1(\widetilde{\mathbf{d}}, \mathbf{u}_1, \widehat{h}^q_{12}(\mathbf{u}_1), \widehat{h}^q_{13}(\mathbf{u}_1)), \forall \theta_{k,1} \subset \Theta_1 \\
\hat{\mathbf{u}}^q_{k,2} &= \underset{\mathbf{u}_2 \in \theta_{k,2}}{\mathrm{argmax}}\, g_2(\widetilde{\mathbf{d}}, \mathbf{u}_2, \widehat{h}^q_{21}(\mathbf{u}_2), \widehat{h}^q_{23}(\mathbf{u}_2)), \forall \theta_{k,2} \subset \Theta_2 \\
\hat{\mathbf{u}}^q_{k,3} &= \underset{\mathbf{u}_3 \in \theta_{k,3}}{\mathrm{argmax}}\, g_3(\widetilde{\mathbf{d}}, \mathbf{u}_3, \widehat{h}^q_{31}(\mathbf{u}_3), \widehat{h}^q_{32}(\mathbf{u}_3)), \forall \theta_{k,3} \subset \Theta_3
\end{aligned} \tag{12}$$

with the corresponding values $\hat{g}^q_{k,1}$, $\hat{g}^q_{k,2}$ and $\hat{g}^q_{k,3}$.

3.3 Complexity Analysis

From the definition of the hyperpower set in (6) it is clear that the number of focal elements increases exponentially with the number of dimensions. Even if one limits the U space to the sole Θ the total number of Focal Elements (FE) for a problem with m uncertain variables, each defined over N_k intervals, is:

$$N_{FE} = \prod_{k=1}^{m} N_k. \tag{13}$$

In terms of coupled and uncoupled uncertain vectors we can write:

$$N_{FE} = \left(\prod_{i=1}^{m_u} \prod_{k=1}^{p_i^u} N^u_{i,k} \right) \left(\prod_{i=1}^{m_c} \prod_{k=1}^{p_i^c} N^c_{i,k} \right). \tag{14}$$

where p_i^u and p_i^c are the number of components of the $i-th$ uncoupled and coupled vector, respectively, and $N^u_{i,k}$ and $N^c_{i,k}$ are the number of intervals of the $k-th$ components of the $i-th$ uncoupled and coupled vector respectively. The total number of focal elements that needs to be explored in the decomposition is instead:

$$N^{Dec}_{FE} = N_s \sum_{i=1}^{m_u} N^u_{FE,i} + \sum_{i=1}^{m_c} N^c_{FE,i} \tag{15}$$

considering the vector of uncertainties ordered as

$$\mathbf{u} = [\underbrace{\mathbf{u}_1, ..., \mathbf{u}_{m_u}}_{\text{uncoupled}}, \underbrace{\mathbf{u}_1, ..., \mathbf{u}_{m_c}}_{\text{coupled}}]$$

where and N_s is the number of samples in the partial belief curves, $N^c_{FE,i} = \prod_{k=1}^{p_i^c} N^c_{i,k}$ and $N^u_{FE,i} = \prod_{k=1}^{p_i^u} N^u_{i,k}$. This means that the computational complexity to calculate the maxima of the function F within the focal elements is polynomial with the number of subsystems and remains exponential for each individual uncoupled or coupled vector.

3.4 Reconstruction

Once all the maxima over the focal elements of the uncoupled variables are available for each sample q one can calculate an approximation of $Bel(F(\mathbf{d},\mathbf{u}) < \nu)$ as follows. From Eq. (12), for each sample q the maximum associated to focal element $\theta_k = \theta_{k_1,1} \times \theta_{k_2,2} \times \theta_{k_3,3}$, for $k = 1, ...N_{FE,1} \cdot N_{FE,2} \cdot N_{FE,3}$, given the condition of positive semidefinition of g_i, is:

$$\max_{(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3)\in\theta_k} F(\tilde{\mathbf{d}}, \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \hat{\mathbf{u}}^q_{12}, \hat{\mathbf{u}}^q_{13}, \hat{\mathbf{u}}^q_{23}) = \hat{g}^q_{k_1,1} + \hat{g}^q_{k_2,2} + \hat{g}^q_{k_3,3} \tag{16}$$

with associated basic probability assignment:

$$bpa^q(\theta_k) = bpa(\theta_{k_1,1}) bpa(\theta_{k_2,2}) bpa(\theta_{k_3,3}) \Delta Bel^q \tag{17}$$

where $\Delta Bel^q = \prod_{ij} \Delta Bel^q_{ij}$ are the contributions of the partial belief curves in (11). In other words, the bpa of each θ_k is the product of all the bpa's of the FE of each uncoupled variable scaled with the product of the belief values of the samples drawn from the partial belief curves(Line 18). The approximation of the belief is then computed as:

$$\widetilde{Bel}(F(\mathbf{d},\mathbf{u}) \le \nu) = \sum_q \sum_k bpa^q(\theta_k) \tag{18}$$

If the decomposition drastically reduces the number of maximisations, the reconstruction still requires an exponential number of multiplications of bpa's. Thus, the computational cost of the reconstruction step would increase exponentially with the number of sub-systems if the full curve was required. In this case the number of times that (17) has to be evaluated would be:

$$N_{\text{evals}} = N_s \prod_{i=1}^{m_u} N^u_{FE,i} \tag{19}$$

If the decomposition is used to evaluate $Bel(F(\mathbf{d},\mathbf{u}) < \nu)$, for a given $\mathbf{d}$ and a single threshold ν, then a partial belief curve could be reconstructed only in a neighborhood of ν at a reduced computational cost.

For a given sample q, consider the vector

$$\hat{\mathbf{g}}^q_i = [\hat{g}^q_{i,1}, ..., \hat{g}^q_{i,N^u_{FE,i}}]^T$$

of all the maxima of a function g_i over all the focal elements $\theta_{k,i}$ and the collection of vectors

$$\Gamma = [\gamma_{qik}] \begin{array}{ll} q & = 1, ..., N_S \\ i & = 1, ..., m_u \\ k & = 1, ..., N^u_{FE,i} \end{array}, \quad \gamma_{qik} = \hat{g}^q_{k,i}$$

organised as in Table 1. The approximated belief curve in Eq. (18) can be computed by taking the sum of the bpa's for every row of Γ and then adding up all the rows.

Now, given ν one can filter out all the components $\hat{g}^q_{k,i}$ of each vector $\hat{\mathbf{g}}^q_i$ that satisfies the relationship:

$$\hat{g}^q_{k,i} + \sum_{i=1}^{N_U} \min_k \hat{\mathbf{g}}^q_i > \nu \tag{20}$$

If condition (20) is applied to every vector in Γ we obtain a new collection Γ_L. Symmetrically we can also construct the collection Γ_R by filtering the vectors in Γ with the following condition:

$$\hat{g}^q_{k,i} + \sum_{i=1}^{N_U} \max_k \hat{\mathbf{g}}^q_i < \nu \tag{21}$$

The computation of $Bel(F(\mathbf{d},\mathbf{u}) \leq \nu)$ is now realised by taking from each row of the two collections Γ_L and Γ_R the ones that contain the least amount of focal elements, i.e. the $\hat{\mathbf{g}}^{\mathbf{q}}_{\mathbf{i}}$ vectors with the lowest number of components, and form the new collection Γ_ν. We can now calculate the approximated belief as in Eq.(18) but using the rows and columns of matrix Γ_ν.

Table 1: information used in the reconstruction step

	sub_1	sub_2	...	sub_i	...	sub_{m_u}
sample$_1$	$\hat{\mathbf{g}}^1_1$	$\hat{\mathbf{g}}^1_2$	...	$\hat{\mathbf{g}}^1_i$	...	$\hat{\mathbf{g}}^1_{m_u}$
sample$_2$	$\hat{\mathbf{g}}^2_1$	$\hat{\mathbf{g}}^2_2$	...	$\hat{\mathbf{g}}^2_i$	...	$\hat{\mathbf{g}}^2_{m_u}$
...	...	...	...	...	...	...
sample$_q$	$\hat{\mathbf{g}}^q_1$	$\hat{\mathbf{g}}^q_2$	...	$\hat{\mathbf{g}}^q_i$	...	$\hat{\mathbf{g}}^q_{m_u}$
...	...	...	...	...	...	...
sample$_{N_s}$	$\hat{\mathbf{g}}^{N_s}_1$	$\hat{\mathbf{g}}^{N_s}_2$	...	$\hat{\mathbf{g}}^{N_s}_i$	...	$\hat{\mathbf{g}}^{N_s}_{m_u}$

The new computational cost of the reconstruction after filtering is:

$$N_{\text{evals}} = \sum_{q=1}^{N_s} \prod_{i=1}^{m_u} \dim(\hat{\mathbf{g}}^{\mathbf{q}}_{\mathbf{i}}), \quad \hat{\mathbf{g}}^{\mathbf{q}}_{\mathbf{i}} \in \Gamma_\nu \tag{22}$$

Figure 2 shows an example of reconstruction, after decomposition, in a neighborhood of $\nu = 300$. The black curve in Figure 2 is the exact belief curve computed for problem in test case 1 in the results section. In this case, the reconstruction without filtering (red curve in Figure 2) requires $N_{\text{evals}} = 2(9 \times 9) = 162$, while the reconstruction after filtering (blue curve in Figure 2) requires $N_{\text{evals}} = 90$.

4 Theoretical Considerations

In this section we prove that the decomposition, as explained in the previous sections, always produces an approximated belief that is lower or equal to the exact one. Suppose that a function F has the following form (for easiness of notation we omit the design variable $\mathbf{d}$ that is fixed during the decomposition and we assume only two

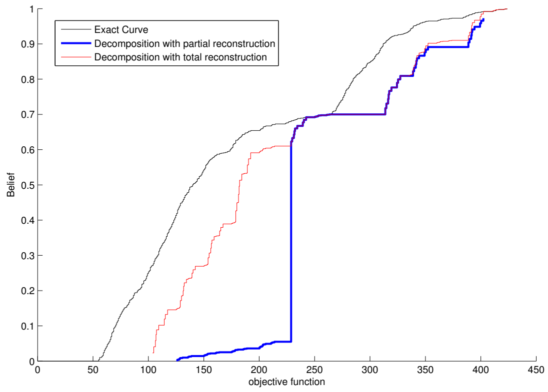

Figure 2: Reconstruction of the Belief curve in the neighborhood of the threshold $\nu = 300$. The objective function is Test case 1 with $\mathbf{u} \in \{[-5,-1] \cup [-3,0] \cup [1,2]\}^6$ and $bpa = [0.3, 0.3, 0.4]^6$; the curve has been reconstructed with two samples in different positions compared with Figure 3.

uncoupled vector and one coupled vector):

$$F(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_{12}) = g_1(\mathbf{u}_1, h_{12}(\mathbf{u}_1,\mathbf{u}_{12})) + g_2(\mathbf{u}_2, h_{21}(\mathbf{u}_2,\mathbf{u}_{12})) \tag{23}$$

with $\mathbf{u}_1 \in \Theta_1$, $\mathbf{u}_2 \in \Theta_2$, $\mathbf{u}_{12} \in \Theta_{12}$ and $\Theta = \Theta_1 \times \Theta_2 \times \Theta_{12}$; and the functions $g_1(\cdot,\cdot)$ and $g_2(\cdot,\cdot)$ are positive definite i.e.

$$\begin{aligned} g_1(\mathbf{u}_1, h_{12}(\mathbf{u}_1,\mathbf{u}_{12})) &\geq 0, \\ g_2(\mathbf{u}_2, h_{21}(\mathbf{u}_2,\mathbf{u}_{12})) &\geq 0 \quad \forall\, (\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_{12}) \in \Theta \end{aligned} \tag{24}$$

If we introduce the spaces $\Omega_{12} \subseteq \Theta_{12}$ and $\Omega = \Theta_1 \times \Theta_2 \times \Omega_{12} \subseteq \Theta$ we can define the maximiser of function F in Ω as:

$$(\underline{\mathbf{u}}_1, \underline{\mathbf{u}}_2, \hat{\mathbf{u}}_{12}) = \underset{\mathbf{u}_{12} \in \Omega_{12}}{\operatorname{argmax}} F(\underline{\mathbf{u}}_1, \underline{\mathbf{u}}_2, \mathbf{u}_{12}) \tag{25}$$

where $\underline{\mathbf{u}}_1$ and $\underline{\mathbf{u}}_2$ are the uncoupled components of vector $\underline{\mathbf{u}}$ solution of problem (3). We can now prove the following lemmas.

Lemma 4.1. *Given a function F as in (23) if the following monotonicity conditions hold true: given $\mathbf{u}^1_{12}$ and $\mathbf{u}^2_{12}$ in Ω_{12}*

$$\begin{array}{rcl} g_1(\mathbf{u}_1, h_{12}(\mathbf{u}_1,\mathbf{u}^1_{12})) & \leq & g_1(\mathbf{u}_1, h_{12}(\mathbf{u}_1,\mathbf{u}^2_{12})) \\ & \updownarrow & \\ h_{12}(\mathbf{u}_1,\mathbf{u}^1_{12}) & \leq & h_{12}(\mathbf{u}_1,\mathbf{u}^2_{12}) \\ & \text{and} & \\ g_2(\mathbf{u}_2, h_{21}(\mathbf{u}_2,\mathbf{u}^1_{12})) & \leq & g_2(\mathbf{u}_2, h_{21}(\mathbf{u}_2,\mathbf{u}^2_{12})) \\ & \updownarrow & \\ h_{21}(\mathbf{u}_2,\mathbf{u}^1_{12}) & \leq & h_{21}(\mathbf{u}_2,\mathbf{u}^2_{12}) \end{array} \tag{26}$$

for all $\mathbf{u}_1 \in \Theta_1$ and $\mathbf{u}_2 \in \Theta_2$. Then

$$\begin{aligned} F(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_{12}) \leq {} & \max_{\mathbf{u}_1 \in \Theta_1} g_1(\mathbf{u}_1, h_{12}(\mathbf{u}_1,\hat{\mathbf{u}}_{12})) \\ & + \max_{\mathbf{u}_2 \in \Theta_2} g_2(\mathbf{u}_2, h_{21}(\mathbf{u}_2,\hat{\mathbf{u}}_{12})) \end{aligned} \tag{27}$$

for all $(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_{12}) \in \Omega$.

Algorithm 1 Decomposition

Initialise
Uncoupled vectors $\mathbf{u}_u = [\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_i, ..., \mathbf{u}_{m_u}]$
Coupled vectors $\mathbf{u}_c = [\mathbf{u}_{12}, \mathbf{u}_{13}, ..., \mathbf{u}_{ij}, ..., \mathbf{u}_{m_c}]$
for a given design $\widetilde{\mathbf{d}}$ **do**
  Compute $(\widetilde{\mathbf{d}}, \underline{\mathbf{u}}_u, \underline{\mathbf{u}}_c)$ = (argmax $F(\widetilde{\mathbf{d}}, \mathbf{u}_u, \mathbf{u}_c)$
  for all $\mathbf{u}_{ij} \in \mathbf{u}_c$ **do**
    for all Focal Elements $\theta_{k,ij} \subseteq \Theta_{ij}$ **do**
      $\widehat{F}_{k,ij} = \max_{\mathbf{u}_{ij} \in \theta_{k,ij}} F(\widetilde{\mathbf{d}}, \underline{\mathbf{u}}_{\mathbf{u}}, \mathbf{u}_{ij})$
      $\widehat{\mathbf{u}}_{k,ij} = \mathrm{argmax}_{\mathbf{u}_{ij} \in \theta_{k,ij}} F$
      Evaluate $bpa(\theta_{k,ij})$
      Evaluate partial Belief curve $Bel(F(\mathbf{u}_{ij}) \leq \nu)$
    end for
    for number of samples **do**
      Evaluate ΔBel^q, $\widehat{\mathbf{u}}_{k,ij}$ and $\widehat{F}_{k,ij}$
    end for
  end for
  for all the combinations of samples **do**
    for all $\mathbf{u}_i \in \mathbf{u}_u$ **do**
      for all Focal Elements $\theta_{k,i} \subseteq \Theta_i$ **do**
        Run $F_{max,k,i} = \max_{\theta_{k,i}} F(\widetilde{\mathbf{d}}, \widehat{\mathbf{u}}_{\mathbf{c}}, \mathbf{u}_i)$
        Evaluate $bpa(\theta_{k,i})$
      end for
    end for
    for all the combinations of Focal Elements $\theta_t \in \Theta_1 \times \Theta_2 \times ... \times \Theta_{m_u}$ **do**
      Evaluate $F_{max,k} \leq \nu$
      Evaluate bpa_k
    end for
    Evaluate the Belief for this sample by constructing collection Γ_ν
  end for
  Add up all belief values for all samples
end for

Proof. Given monotonicity condition (26) and the fact that functions $g_1(\cdot,\cdot)$ and $g_2(\cdot,\cdot)$ are positive semidefined, we have:

$$\begin{aligned} F(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_{12}) &= g_1(\mathbf{u}_1, h_{12}(\mathbf{u}_1, \mathbf{u}_{12})) + g_2(\mathbf{u}_2, h_{21}(\mathbf{u}_2, \mathbf{u}_{12})) \\ &\leq g_1(\mathbf{u}_1, h_{12}(\mathbf{u}_1, \hat{\mathbf{u}}_{12})) + g_2(\mathbf{u}_2, h_{21}(\mathbf{u}_2, \hat{\mathbf{u}}_{12})) \\ &\leq \max_{\mathbf{u}_1 \in \Theta_1} g_1(\mathbf{u}_1, h_{12}(\mathbf{u}_1, \hat{\mathbf{u}}_{12})) + \\ &\quad + \max_{\mathbf{u}_2 \in \Theta_2} g_2(\mathbf{u}_2, h_{21}(\mathbf{u}_2, \hat{\mathbf{u}}_{12})) \end{aligned} \tag{28}$$

for all $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_{12}) \in \Omega$.

□

Lemma 4.2. *Given a function F as in (23), if monotonicity condition (26) holds true then there is at least one $\hat{\mathbf{u}}_{12} \in \theta_{12} \subset \Omega_{12}$ such that:*

$$\max_{(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_{12}) \in \Omega} F(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_{12}) = \max_{\mathbf{u}_1 \in \Theta_1} g_1(\mathbf{u}_1, h_{12}(\mathbf{u}_1, \hat{\mathbf{u}}_{12})) + \max_{\mathbf{u}_2 \in \Theta_2} g_2(\mathbf{u}_2, h_{21}(\mathbf{u}_2, \hat{\mathbf{u}}_{12})) \tag{29}$$

Proof. From Lemma 4.1, since the inequality holds $\forall (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_{12}) \in \Omega$ it holds also for the value that attains the maximum:

$$\max_{(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_{12}) \in \Omega} F(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_{12}) \leq \max_{\mathbf{u}_1 \in \Theta_1} g_1(\mathbf{u}_1, h_{12}(\mathbf{u}_1, \hat{\mathbf{u}}_{12})) + \max_{\mathbf{u}_2 \in \Theta_2} g_2(\mathbf{u}_2, h_{21}(\mathbf{u}_2, \hat{\mathbf{u}}_{12})) \tag{30}$$

Let us assume now that the solution $(\mathbf{u}_1^*, \mathbf{u}_2^*, \mathbf{u}_{12}^*) := \mathrm{argmax}_{(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_{12}) \in \Omega} F(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_{12})$ is such that:

$$F(\mathbf{u}_1^*, \mathbf{u}_2^*, \mathbf{u}_{12}^*) < \max_{\mathbf{u}_1 \in \Theta_1} g_1(\mathbf{u}_1, h_{12}(\mathbf{u}_1, \hat{\mathbf{u}}_{12})) + \max_{\mathbf{u}_2 \in \Theta_2} g_2(\mathbf{u}_2, h_{21}(\mathbf{u}_2, \hat{\mathbf{u}}_{12})) \tag{31}$$

In this case solution $(\mathbf{u}_1^*, \mathbf{u}_2^*, \mathbf{u}_{12}^*)$ is not a global maximiser. As a consequence if solution $(\mathbf{u}_1^*, \mathbf{u}_2^*, \mathbf{u}_{12}^*)$ is a global maximum of F then it must be that:

$$\max_{(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_{12}) \in \Omega} F(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_{12}) = \max_{\mathbf{u}_1 \in \Theta_1} g_1(\mathbf{u}_1, h_{12}(\mathbf{u}_1, \hat{\mathbf{u}}_{12})) + \max_{\mathbf{u}_2 \in \Theta_2} g_2(\mathbf{u}_2, h_{21}(\mathbf{u}_2, \hat{\mathbf{u}}_{12})) \tag{32}$$

□

Lemma 4.1 and 4.2 allow us to demonstrate the following fundamental theorem on the accuracy of the decomposition.

Theorem 4.3. *If the decomposition is applied to calculate an approximation $\widetilde{Bel}(F(\mathbf{u}) < \nu)$ of $Bel(F(\mathbf{u}) < \nu)$ and monotonicity condition (26) holds true, then:*

$$\widetilde{Bel}(F(\mathbf{u}_1, \mathbf{u}_2, \hat{\mathbf{u}}_{12}) \leq \nu) \leq Bel(F(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_{12}) \leq \nu) \tag{33}$$

for $\mathbf{u}_1 \in \Theta_1$, $\mathbf{u}_2 \in \Theta_2$, $\mathbf{u}_{12} \in \Omega_{12}$, $\hat{\mathbf{u}}_{12} \in \hat{\Omega}_{12}$ and $\Omega_{12} \subset \hat{\Omega}_{12}$. Furthermore,

$$\widetilde{Bel}(F(\mathbf{u}_1, \mathbf{u}_2, \hat{\mathbf{u}}_{12}) \leq \nu) = Bel(F(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_{12}) \leq \nu) \tag{34}$$

when $\Omega_{12} = \hat{\Omega}_{12}$.

Proof. For a given ν consider the corresponding set of focal elements

$$\Omega_\nu = \bigcup_t \{\theta_t : \max_{u_1, u_2, u_{12} \in \theta_t} F(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_{12}) < \nu\} \tag{35}$$

where for every t it exists a set of indexes $\{i, j, k\}$ such that

$$\theta_t = \theta_{1,i} \times \theta_{2,j} \times \theta_{12,k} \tag{36}$$

then the exact cumulative belief function is:

$$Bel(F(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_{12}) \le \nu) = \sum_{t \to \{i,j,k\}} bpa(\theta_{1,i})bpa(\theta_{2,j})bpa(\theta_{12,k}) \quad (37)$$

Consider now the approximated cumulative belief function $\widetilde{Bel}^q(F(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_{12}) \le \nu)$ constructed as follows:

$$\widetilde{Bel}^q(F(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_{12}) \le \nu) = \sum_i \sum_j bpa(\theta_{1i})bpa(\theta_{2,j})Bel^q(\Omega_{12}) \quad (38)$$

for all $\theta_{1,i} \in \Theta_1$ and $\theta_{2,j} \in \Theta_2$ such that:

$$\max_{\mathbf{u}_1 \in \theta_{1,i}} g_1(\mathbf{u}_1, \hat{\mathbf{u}}_{12}^q) + \max_{\mathbf{u}_2 \in \theta_{2,j}} g_2(,\mathbf{u}_2, \hat{\mathbf{u}}_{12}^q) < \nu \quad (39)$$

with

$$\hat{\mathbf{u}}_{12}^q = \underset{\mathbf{u}_{12} \in \Omega_{12}^q}{\operatorname{argmax}} F(\underline{\mathbf{u}}_1, \underline{\mathbf{u}}_2, \mathbf{u}_{12}) \quad (40)$$

Note that for a single sample q the ΔBel^q in (17) is computed from $Bel = 0$, therefore, $\Delta Bel^q = Bel^q$. From Lemma 4.1 we can say that:

$$\begin{aligned}&\max_{(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_{12}) \in \theta_t} F(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_{12}) < \\ &\max_{\mathbf{u}_1 \in \theta_{1,i}} g_1(\mathbf{u}_1, \hat{\mathbf{u}}_{12}^q) + \max_{\mathbf{u}_2 \in \theta_{2,j}} g_2(\mathbf{u}_2, \hat{\mathbf{u}}_{12}^q) \\ &\forall q | \Omega_\nu \subset \Omega_q\end{aligned} \quad (41)$$

which means that by construction

$$\widetilde{Bel}^q(F(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_{12}) \le \nu) \le Bel(F(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_{12}) \le \nu).$$

From Lemma 4.2 we know that for $\Omega_\nu = \Omega_q$:

$$\begin{aligned}&\max_{(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_{12}) \in \Omega_\nu} F(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_{12}) = \\ &\max_{\mathbf{u}_1 \in \Theta_1} g_1(\mathbf{u}_1, \hat{\mathbf{u}}_{12}^q) + \max_{\mathbf{u}_2 \in \Theta_2} g_2(\mathbf{u}_2, \hat{\mathbf{u}}_{12}^q)\end{aligned} \quad (42)$$

Therefore, comparing Eq.(38) with Eq.(37) we can say that $\widetilde{Bel}^q(F(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_{12}) \le \nu) = Bel(F(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_{12}) \le \nu)$. □

Eq. (41) suggests a possible definition of monotonicity for multivariate functions. Let $P(\Theta)$ be the set of all the subsets of Θ. If Ω_1 and Ω_2 are two elements of $P(\Theta)$ we say that F is monotonically increasing with respect to $\mathbf{u} \in P(\Theta)$ if:

$$\max_{\mathbf{u} \in \Omega_1} F(\mathbf{u}) \le \max_{\mathbf{u} \in \Omega_2} F(\mathbf{u}) \Leftrightarrow \Omega_1 \subseteq \Omega_2 \quad (43)$$

5 Surrogate Approach

The decomposition drastically reduces the number of optimisations required to calculate the cumulative belief function. However, the number of focal elements for each coupled or uncoupled vector can be considerable. Thus, in this paper we propose the use of a surrogate of the space of the maxima over the indexes of the focal elements to further reduce the number of maximisations required to compute the belief curve.

To be more specific, if j_i is the index identifying the $j-th$ interval along dimension i, assuming that all intervals are adjacent and the frame of discernment is limited to Θ, the idea is to construct an approximation to the function $F^*(j_1,...,j_i,...,j_m) : \mathbb{N} \to \mathbb{R}$ such that, for each focal element θ_k defined by the product of a particular combination of intervals identified by the vector of indexes $\mathbf{J}_k = [j_1, j_2, ... j_i, ... j_m]_k^T$, the value $F^*(\mathbf{J}_k)$ is the maximum of $F(\mathbf{u})$ with $\mathbf{u} \in \theta_k$.

Given $\tilde{\mathbf{d}} \in D$, we will use a Kriging model[Tardioli et al(2015)Tardioli, Kubicek, Vasile, Minisci, and Riccardi] to interpolate some of the maxima F^* evaluated for a limited set index vectors $\mathbf{J}_k$, and to predict the value of the maxima at other locations (other combinations of intervals). Once an estimation of the belief is computed with the surrogate model a further refinement can be obtained via decomposition or full belief calculation.

The surrogate of the function $F*$ could be constructed also over the U space. However, in this case, in order to sample the surrogate one would need to generate sample vectors $\mathbf{u}$ in such a way that each vector falls in only one focal element. Thus, one would need, anyway, to first take a sample of indexes and then generate a $\mathbf{u}$ vector within the focal element corresponding to the sampled index set.

Note that in this specific case, as the number of intervals per dimension tends to infinity, the number of focal elements also tends to infinity and the distribution of bpa's approaches a continues density function.

6 Robust Design Trade-off Curve

Once the cumulative belief value is computable for each design $\mathbf{d}$ and each ν, one can maximise the belief associated to ν by selecting an optimal $\mathbf{d}$. This can be formulated as the following bi-objective problem:

$$\begin{aligned}&\min \nu \\ &\max_d Bel(F < \nu)\end{aligned} \quad (44)$$

The value of ν for which Bel is maximal, can be obtained by solving a deterministic global min-max problem:

$$\nu_{max} = \min_{\mathbf{d}} \max_{\mathbf{u}} F(\mathbf{d}, \mathbf{u}) \quad (45)$$

Likewise the best value of ν for which $Pl = 0$ can be obtained by solving the deterministic global min-min problem:

$$\nu_{min} = \min_{\mathbf{d}} \min_{\mathbf{u}} F(\mathbf{d}, \mathbf{u}) \quad (46)$$

Once the min-max and min-min values are available one can use the decomposition technique to estimate Bel for every $\nu \in [\nu_{min}, \nu_{max}]$ and optimise the whole Bel curve following Algorithm 2. Note that the result is a set of $\mathbf{d}$ values in which each $\mathbf{d}$ provides the maximum belief only for a given ν. This means that the belief curve corresponding to the $\mathbf{d}$ associated to a particular ν can be suboptimal for other ν's.

7 Results

The Decomposition and the Surrogate algorithms are here applied to five test cases, four of which use synthetic functions and one considers a real space system design

Algorithm 2 Robust Trade-off Curve

```
Initialisation
Compute ν_min = min min and ν_max = min max
for each ν ∈ [ν_min, ν_max] do
    Compute max_d Bel~(F ≤ ν)
end for
```

problem. The first four cases are testing the decomposition algorithm and its ability to deliver a good approximation to the true belief curve at a fraction of the computational cost. The fifth case will test the use of the surrogate model and the optimisation of the belief curve for different ν's.

7.1 Decomposition Test Set

In test cases 1 and 2 each partial belief curves was sampled the same number of times $N_s = cost^{m_c}$.

On the contrary, in test cases 3 and 4, different partial Belief curves were sampled differently so that the number of samples is:

$$N_s = \prod_{i=1}^{m_c} N_{s,i}$$

Table 2 reports the number of intervals per dimension, in the uncertainty space, the range of the **d** vector and the bpa's associated to each interval, for cases 1,2 and 3. Case 1 in particular was solved for an increasing number of intervals per dimension. In all tests, the uncertainty space is $U = \Theta$ and uncertain variables are uncorrelated.

Table 2: Design, uncertain parameters and bpa structure for test cases 1, 2 and 3.

#	Parameters
1	$\mathbf{d} \in [-5,5]^2$ $\mathbf{u} \in \{[-5,-1],[-3,0],[1,2]\}^6$ $\mathbf{u} \in \{[-5,-1],[-3,0],[1,2],[3,5]\}^6$ $\mathbf{u} \in \{[-5,-1],[-3,0],[1,2],[2.5,4],[3,5]\}^6$ $\mathbf{u} \in \{[-5,-1],[-3,0],[1,2],[2.1,3],[3.1,4.5],[4,5]\}^6$ $\mathbf{u} \in \{[-5,-1],[-3,0],[1,2],[2.1,3],$ $[3.1,4.5],[4,5],[3,7]\}^6$
2	$\mathbf{d} \in [0,3]^{13}$ $\mathbf{u} \in \{[-5,-4],[-3,0],[-1,3]\}^{13}$ bpa = $[0.1,0.25,0.65]^{13}$
3	$\mathbf{d} \in [0,3]^8$ $\mathbf{u} \in \{[-5,-4],[-3,0],[-1,3]\}^8$ bpa = $[0.1,0.25,0.65]^8$

7.1.1 Decomposition: test case 1

In test case 1, the cost function F is composed of two sub-functions g_1 and g_2:

$$\begin{cases} F = g_1 + g_2 \\ g_1 = 10u_1^2 + |u_2|u_5^2 + \frac{u_6^4}{100} + d_1|d_2| \\ g_2 = |u_3| + u_4^2\frac{|u_5|}{10} + u_6^2 + |d_1| \end{cases}$$

The uncertain vector **u** is:

$$\begin{cases} \mathbf{u} = (\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_{12}}) & \\ dim(\mathbf{u_1}) = 2 & \mathbf{u_1} = (u_1, u_2) \\ dim(\mathbf{u_2}) = 2 & \mathbf{u_2} = (u_3, u_4) \\ dim(\mathbf{u_{12}}) = 2 & \mathbf{u_{12}} = (u_5, u_6) \end{cases}$$

where $\mathbf{u}_{12}$ is the vector of coupled variables and $\mathbf{u}_1$ and $\mathbf{u}_2$ are the vectors of uncoupled variables. Different belief curves were computed for an increasing number of intervals per dimension. Tab 3 shows the computational cost of the decomposition compared to an exact calculation of the curve. The first column is the number of intervals per dimension, the second column the number of maximisation required for an exact calculation of the belief, while the third column is the number of maximisation used in the decomposition. The last column is the computational time of the decomposition with respect to the exact calculation. The gain in computational time increases as the number of intervals increases.

Tab 4 and Fig 3 show, for a fixed number of interval $\mathbf{u} \in \{[-5,-1] \cup [-3,0] \cup [1,2]\}^6$, the convergence of the curve computed with the decomposition for an increasing number of samples. The last column in Tab 4 is the relative computational cost. To be noted that for 6 samples the approximated curve is almost identical to the exact one but the computational cost is only 16%.

Table 3: Test case 1; results with 4 samples in the partial Belief curve.

$N_{intervals}$	$N_{opt-exact}$	$N_{opt-Dec}$	cpu cost (%)
3	729	81	11.11
4	4096	144	3.52
5	15625	225	1.44
6	46656	324	0.69
7	117649	729	0.61

Table 4: Test case 1; results with 3 intervals for each dimension of **u** and different number of samples.

$N_{samples}^{partial}$	$N_{samples}^{total}$	$N_{opt-exact}$	$N_{opt-Dec}$	cpu cost (%)
1	1	729	27	3.7
2	2	729	45	6.17
3	3	729	63	8.64
4	4	729	81	11.11
5	5	729	99	13.58
6	6	729	117	16.05

7.1.2 Decomposition: test case 2

In test case 2, the function F is composed of four sub-functions:

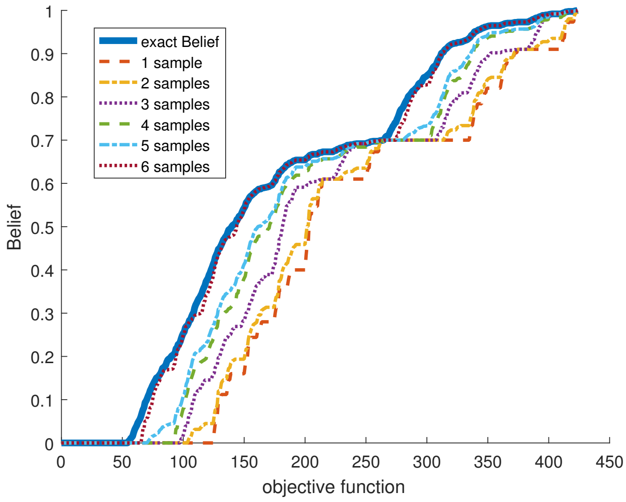

Figure 3: Test case 1, convergence with $\mathbf{u} \in \{[-5,-1] \cup [-3,0] \cup [1,2]\}^6$ and $bpa = [0.3, 0.3, 0.4]^6$

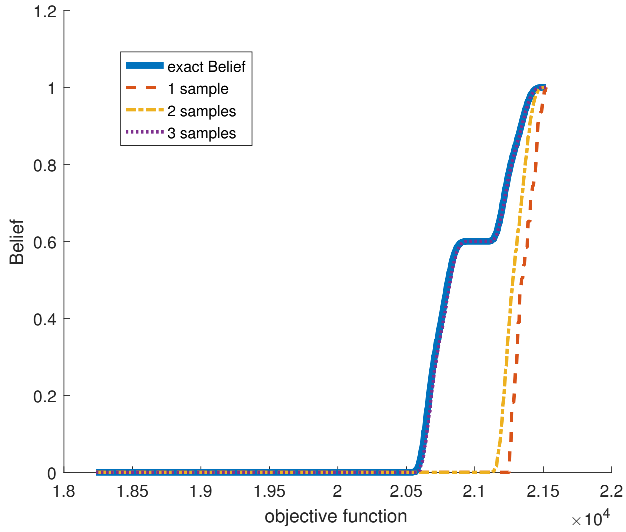

Figure 4: Test case 2, convergence with $bpa = [0.1, 0.25, 0.65]^{13}$

$$\begin{cases} F = g_1 + g_2 + g_3 + g_4 \\ g_1 = \sum_{i=1}^{2} (d_i - u_i)^2 + \frac{1}{2} \sum_{i=6}^{9} (d_i - u_i)^2 \\ g_2 = (d_3 - u_3)^2 + \frac{1}{2} \sum_{i=6}^{7} (d_i - u_i)^2 + \frac{1}{2} \sum_{i=10}^{11} (d_i - u_i)^2 \\ g_3 = (d_4 - u_4)^2 + \frac{(d_8 - u_8)^2}{2} + \frac{(d_{10} - u_{10})^2}{2} + \frac{1}{2} \sum_{i=12}^{13} (d_i - u_i)^2 \\ g_4 = (d_5 - u_5)^2 + \frac{1}{2} (d_9 - u_9)^2 + \frac{1}{2} \sum_{i=11}^{13} (d_i - u_i)^2 \end{cases}$$

where the uncertain vector $\mathbf{u}$ is composed of four uncoupled vectors: $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ and $\mathbf{u}_4$, and six coupled vectors, $\mathbf{u}_{12}$, $\mathbf{u}_{13}$, $\mathbf{u}_{14}$, $\mathbf{u}_{23}$, $\mathbf{u}_{24}$ and $\mathbf{u}_{34}$. The results are shown in Fig 4 and Tab 5 In this case only 3 samples are sufficient to achieve almost the exact value of the belief, though with a computational cost of 0.82% of the exact one.

$$\begin{cases} \mathbf{u} = (\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}, \mathbf{u_4}, & \\ \quad \mathbf{u_{12}}, \mathbf{u_{13}}, \mathbf{u_{14}}, \mathbf{u_{23}}, \mathbf{u_{24}}, \mathbf{u_{34}}) & \\ dim(\mathbf{u_1}) = 2 & \mathbf{u_1} = (u_1, u_2) \\ dim(\mathbf{u_2}) = 1 & \mathbf{u_2} = (u_3) \\ dim(\mathbf{u_3}) = 1 & \mathbf{u_3} = (u_4) \\ dim(\mathbf{u_4}) = 1 & \mathbf{u_4} = (u_5) \\ dim(\mathbf{u_{12}}) = 2 & \mathbf{u_{12}} = (u_6, u_7) \\ dim(\mathbf{u_{13}}) = 1 & \mathbf{u_{13}} = (u_8) \\ dim(\mathbf{u_{14}}) = 1 & \mathbf{u_{14}} = (u_9) \\ dim(\mathbf{u_{23}}) = 1 & \mathbf{u_{23}} = (u_{10}) \\ dim(\mathbf{u_{24}}) = 1 & \mathbf{u_{24}} = (u_{11}) \\ dim(\mathbf{u_{34}}) = 2 & \mathbf{u_{34}} = (u_{12}, u_{13}) \end{cases}$$

Table 5: Test case 2 with different number of samplings.

$N^{partial}_{samples}$	$N^{total}_{samples}$	$N_{opt-exact}$	$N_{opt-Dec}$	cpu cost (%)
1	1	1594323	48	$3e-3$
2	64	1594323	1182	$7.4e-2$
3	729	1594323	13152	0.82
4	4096	1594323	73758	4.6

7.1.3 Decomposition: test case 3

The function F is here composed of three sub-functions:

$$\begin{cases} F = g_1 + g_2 + g_3 \\ g_1 = (d_1 - u_1)^2 + \frac{1}{2} \sum_{i=4}^{7} (d_i - u_i)^2 \\ g_2 = (d_2 - u_2)^2 + \frac{1}{2} \sum_{i=4}^{5} (d_i - u_i)^2 + \frac{1}{2} (d_8 - u_8)^2 \\ g_3 = (d_3 - u_3)^2 + \frac{1}{2} \sum_{i=6}^{8} (d_i - u_i)^2 \end{cases}$$

where the uncertain vector $\mathbf{u}$ is composed of three uncoupled vectors, $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$, and three coupled vectors, $\mathbf{u}_{12}$, $\mathbf{u}_{13}$ and $\mathbf{u}_{23}$. The convergence and computational cost results are reported in Fig 5 and Tab 6. Fig 6, instead, shows how the choice of the samples can affect the accuracy of the computation. The blue curve in the figure is the exact belief. The red dashed curve was generated with 18 samples taken in a neighborhood of the left extreme of the partial belief curves, while the yellow dotted curve was generated using samples from a neighborhood of the right most extreme of

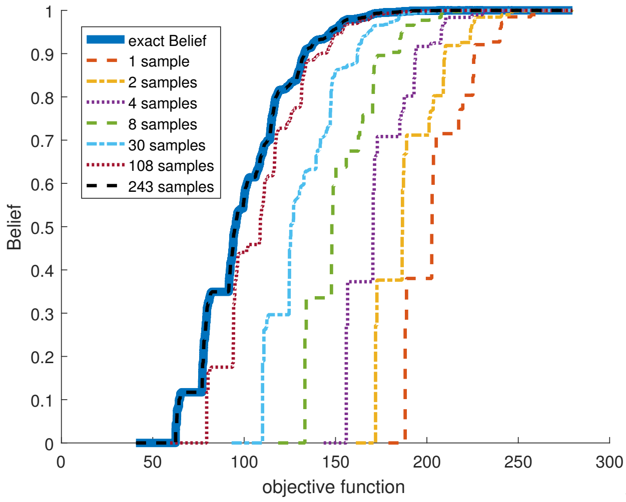

Figure 5: Test case 3, convergence with $bpa = [0.1, 0.25, 0.65]^8$

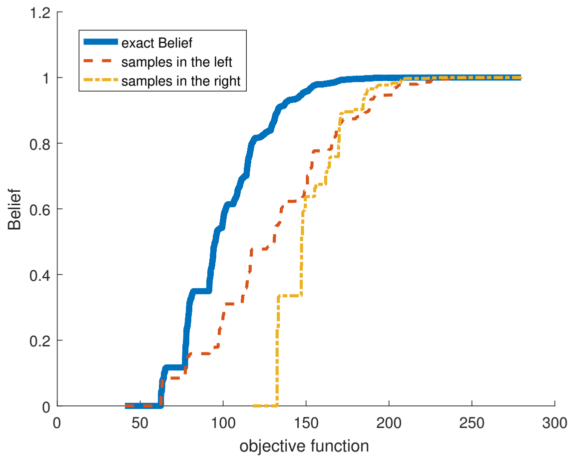

Figure 6: Test case 3. Belief curves generated with a fixed number of samples taken in two different parts of the partial belief curves.

the partial curves.

$$\begin{cases} \mathbf{u} = (\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}, \mathbf{u_{12}}, \mathbf{u_{13}}, \mathbf{u_{23}}) & \\ dim(\mathbf{u}_1) = 1 & \mathbf{u_1} = (u_1) \\ dim(\mathbf{u}_2) = 1 & \mathbf{u_2} = (u_2) \\ dim(\mathbf{u}_3) = 1 & \mathbf{u_3} = (u_3) \\ dim(\mathbf{u}_{12}) = 2 & \mathbf{u_{12}} = (u_4, u_5) \\ dim(\mathbf{u_{13}}) = 2 & \mathbf{u_{13}} = (u_6, u_7) \\ dim(\mathbf{u}_{23}) = 1 & \mathbf{u_{23}} = (u_8) \end{cases}$$

Table 6: Test case 3 with different number of samplings.

$N_{samples}^{total}$	$N_{opt-exact}$	$N_{opt-Dec}$	cpu cost (%)
1	6561	30	0.46
2	6561	39	0.59
4	6561	57	0.87
8	6561	93	1.41
30	6561	291	4.44
108	6561	993	15.13
243	6561	2208	33.7

7.1.4 Decomposition: test case 4 Spacecraft

The Decomposition technique is here applied to the minimisation of the mass of a spacecraft, as shown in Figure 7. The system is composed of three sub-systems, whose model can be found in this paper.[Alicino and Vasile(2014a)] The mass of the Attitude and Orbit Control System (AOCS), M_{AOCS}, is the sum of the mass of the reaction wheel M_{rw} and magneto-torque M_{mag}. The same two components give the total power consumption P_{AOCS}:

$$\begin{aligned} M_{AOCS} = M_{rw} + M_{mag} \\ P_{AOCS} = P_{rw} + P_{mag} \end{aligned} \tag{47}$$

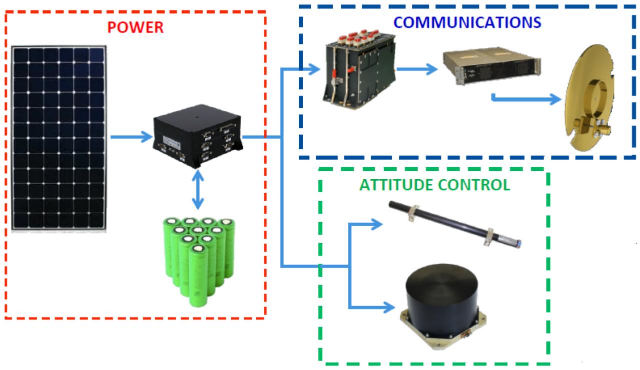

Figure 7: Schematic of spacecraft sub-systems

The Telemetry and Telecommand System (TTS) is composed of an antenna, with mass M_{ant}, a set of amplified transponders, with mass M_{amp}, and a radio frequency distribution network (RFDN), with mass M_{rfdn}. The total mass and power requirement of the TTS is:

$$\begin{aligned} M_{TTC} = M_{ant} + M_{amp} + M_{rfdn} \\ P_{TTC} = P_{amp} \end{aligned} \tag{48}$$

The AOCS and TTS submit their power requirements to the Electrical Power System (EPS). The EPS is composed of a solar array, a battery pack, a power conditioning and distribution unit (PCDU). The total mass of the power system is:

$$M_{EPS} = M_{sa} + M_{batt} + M_{pcdu} \tag{49}$$

The mass of each element of the power system is a monotonic function of the power requirement. The power requirement is the sum of the P_{AOCS} and P_{TTS}. Therefore, the variables defining the power demand of TTS and AOCS are coupled variables and their effect manifests through P_{TTS} and P_{AOCS} respectively.

In this model, the design vector consists of 10 components (dim($\mathbf{d}$)=10), while the uncertain vector has 16 components (dim($\mathbf{u}$)=16), out of which, 11 uncertain components influence one and only one of the functions, thus they are collected in the uncoupled vector:

$$\begin{cases} \mathbf{u}_{un-c} = (\mathbf{u}_{AOCS}, \mathbf{u}_{TTC}, \mathbf{u}_{EPS}) \\ dim(\mathbf{u}_{AOCS}) = 4 \\ dim(\mathbf{u}_{TTC}) = 2 \\ dim(\mathbf{u}_{EPS}) = 5 \end{cases}$$

The other 5 uncertain variables belong to the coupled vector:

$$\begin{cases} \mathbf{u}_{c} = (\mathbf{u}_{AOCS \to TTC}, \mathbf{u}_{AOCS \to EPS}, \mathbf{u}_{TTC \to EPS}) \\ dim(\mathbf{u}_{AOCS \to TTC}) = 0 \\ dim(\mathbf{u}_{AOCS \to EPS}) = 2 \\ dim(\mathbf{u}_{TTC \to EPS}) = 3 \end{cases}$$

As one could see in Figure 8, there is a unidirectional flow of information: AOCS influences EPS but on the other hand EPS's variables are not input to the AOCS; similarly for TTC and EPS; Figure 8 explains that AOCS and TTC influence EPS respectively with their power requirement.

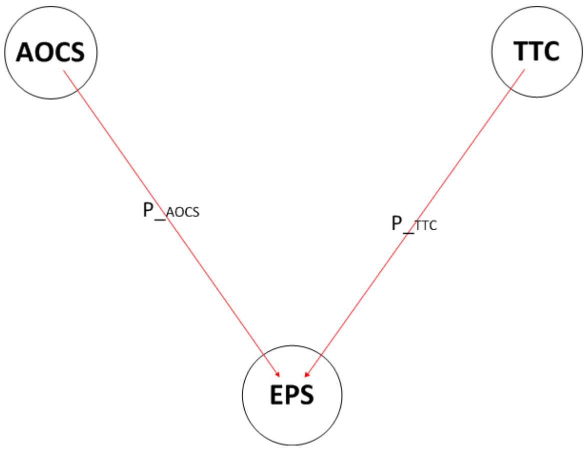

Figure 8: Decomposition of the spacecraft system

The results obtained by applying the decomposition algorithm are shown in Figure 9. The number of optimisations required to generate the approximated curve can be estimated to be:

$$N_{optimisation} = 12 + 52N_s \tag{50}$$

The computational cost for each approximation can be found in Table 7.

7.2 *Surrogate Approach and Robust Trade-off*

The surrogate method is here applied to test case 1. First we computed the error in the representation of the space of the

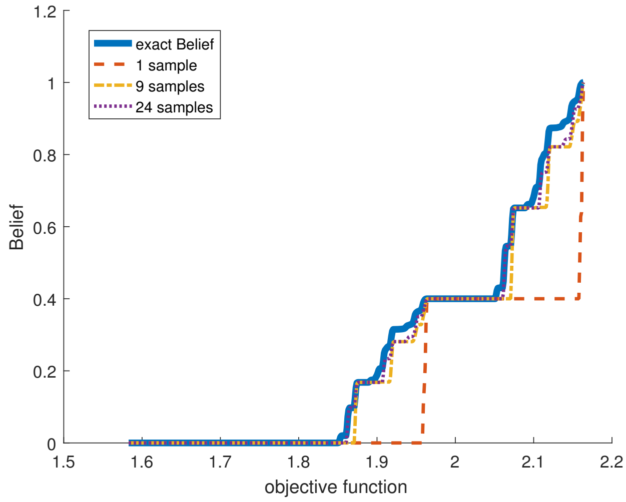

Figure 9: Convergence for the Spacecraft's Belief curve with different numbers of samples (1, 9 and 24).

Table 7: Test case 4

$N^{total}_{samples}$	$N_{opt-exact}$	$N_{opt-Dec}$	cpu cost (%)
1	65536	64	$9.7656e-04$
4	65536	220	0.34
9	65536	480	0.73
16	65536	844	1.29
20	65536	1052	1.61
24	65536	1260	1.92

maxima for different numbers of intervals per dimension (see Tab 8). As expected, as the number of intervals per dimension increases, the space of the maxima tends to a continuous function and the representation with the surrogate becomes more and more accurate.

Then we addressed the solution of problem (44). The Pareto front in Fig 10 is the solution of problem (44). Each point along the curve was optimised using the surrogate to calculate the belief. Figure 11 show the design components of the elements in the Pareto Front.

Then for the $\mathbf{d}$ vector that corresponds to the circled point, the belief curve was recomputed with the decomposition approach. Fig 12 compares the exact curve with the ones obtained with the surrogate and decomposition algorithms. The legend in the figure includes the number of function evaluations for the three cases. As one can see the full calculation, even for such a small dimensional problem with only 2 sub-functions, require a few million function evaluations. The decomposition reduces this number to few hundred thousands and the surrogate to a few thousands.Fig 13 compares the results obtained with different number of maxima used to build the surrogate. As the number of sampled maxima increases the belief curve calculated with the surrogate converges to the exact one. The function

Table 8: Test case 1. Surrogate estimation error for an increasing number of focal elements.

maxima evaluation	50	100	150	200	400
$N_{FE}^{problem}$			729		
maxima evaluation (%)	6.8	13.7	20.5	27.4	54.8
error (%)	22.7	9.9	6.9	3.7	1.1
$N_{FE}^{problem}$			1728		
maxima evaluation (%)	2.9	5.7	8.7	11.6	23.
error (%)	22.6	10.9	6.5	5.1	2.9
$N_{FE}^{problem}$			15625		
maxima evaluation (%)	3.2e-3	6.4e-3	9.6e-3	1.3	2.6
error (%)	21.2	12	10.7	8.2	4.4

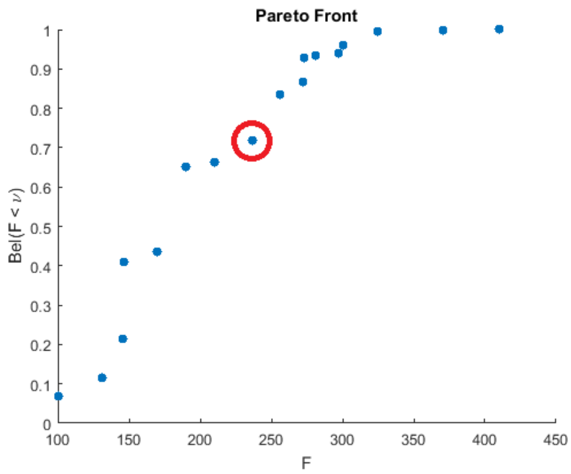

Figure 10: Robust Pareto Front with Surrogate Model (test case 1)

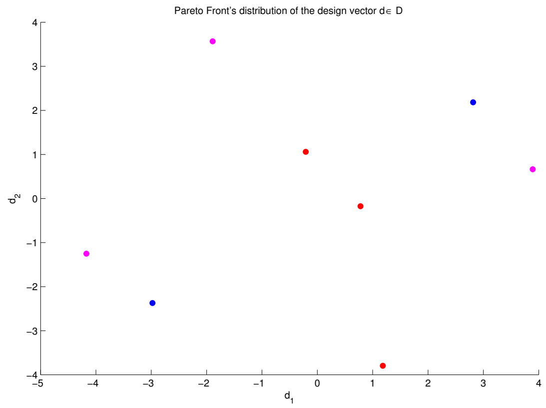

Figure 11: Design components of the surrogate Pareto Front: blue component correspond to an objective function $F < 200$, magenta ones to $200 < F < 300$ and red ones to $F > 300$.

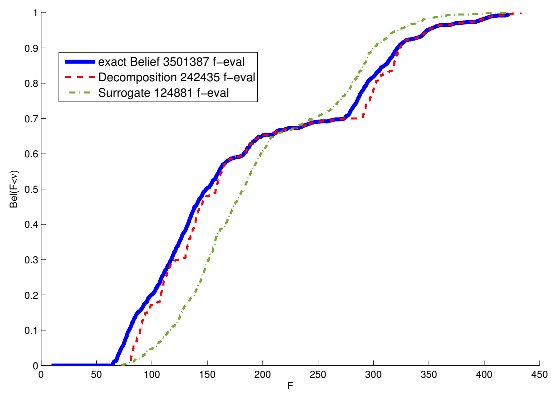

Figure 12: Comparison between the exact Belief curve and the reconstruction with the Decomposition and the Surrogate method. The legend includes the number of function evaluations in the three cases.

evaluations also increases to a value that is anyway one order of magnitude lower than for a full calculation.

8 Conclusions

In this paper we introduced the concept of Evidence Network Models to represent complex engineering systems composed of a number of interconnected sub-systems. The uncertainty associated to each of the sub-systems and their interconnections is modeled with evidence theory. The calculation of the belief in the total value of the ENM is shown to be exponentially complex in the general case. Therefore, a decomposition algorithm is introduced to obtain an approximation in polynomial time. Furthermore, a surrogate-based approach is proposed to further reduce the computational cost when the belief needs to be maximised with respect to the decision (or design) variables.

The methodology is applied to a number of test cases, proving that, under suitable assumptions, a good approximation can be obtained at a fraction of the computational cost of an exact calculation. The paper

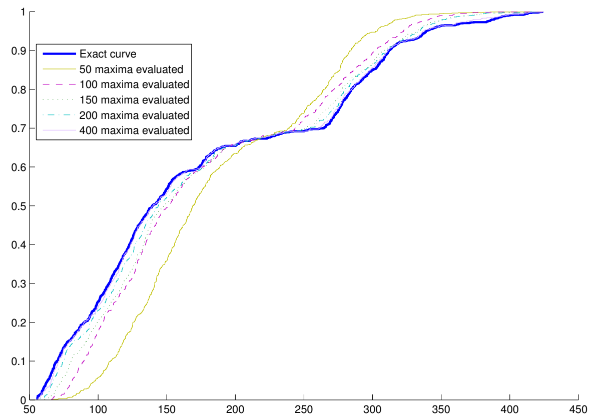

Figure 13: Convergence of the Surrogate method applied to Test case 1.

proposed also one theorem and two lemmas that proof that the approximated belief is always lower or equal to the exact one. This is a very important property as it provides a conservative expectation in the occurrence of an event or the truth of a proposition.

It was also shown that the method proposed in this paper allows for the fast estimation of the total belief of the network at a cost that is polynomial with the number of subsystems. This property is very important as it allows for an increases of the size and complexity of the system while maintaining the computation of the belief affordable. More work is required to study the behaviour of the ENM and estimation algorithms for different topologies and considering the full hyper-power set.

9 Acknowledgements

This work is partially supported by ESTECO Spa. and Surrey Satellite Technologies Ltd. through the project *Robust Design Optimization of Space Systems* (European Space Agency - Innovation Triangle Initiative).

References

[Alicino and Vasile(2014a)] Alicino S, Vasile M (2014a) Evidence-based preliminary design of spacecraft. In: SECESA 2014

[Alicino and Vasile(2014b)] Alicino S, Vasile M (2014b) An evolutionary approach to the solution of multi-objective min-max problems in evidence-based robust optimization. In: 2014 IEEE Congress on Evolutionary Computation (CEC2014)

[Ortega and Vasile(2017)] Ortega C, Vasile M (2017) New heuristics for multi-objective worst-case optimization in evidence-based robust design. In: 2017 IEEE Congress on Evolutionary Computation (CEC2017)

[Shafer(1976a)] Shafer G (1976a) A Mathematical Theory of Evidence. (English). Princeton University Press

[Shafer(1976b)] Shafer G (1976b) A Mathematical Theory of Evidence. Princeton University Press

[Tardioli et al(2015)Tardioli, Kubicek, Vasile, Minisci, and Riccardi] Tardioli C, Kubicek M, Vasile M, Minisci E, Riccardi A (2015) Comparison of non-intrusive approaches to uncertainty propagation in orbital mechanics. AAS Astrodynamics Specialists Conference, Vail, Colorado, USA AAS 15-545

[Vasile(2014)] Vasile M (2014) On the solution of min-max problems in robust optimization. In: The EVOLVE 2014 International Conference