Joslyn, Cliff A and Helton, Jon C: (2002) "Bounds on Plausibility and Belief of Functionally Propagated Random Sets", in: Proc 2002 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS-FLINT 2002), pp. 412-417

Bounds on Belief and Plausibility of Functionally Propagated Random Sets

Cliff Joslyn and Jon C. Helton

Abstract— We are interested in improving risk and reliability analysis of complex systems where our knowledge of system performance is provided by large simulation codes, and where moreover input parameters are known only imprecisely. Such imprecision lends itself to interval representations of parameter values, and thence to quantifying our uncertinay through Dempster-Shafer or Probability Bounds representations on the input space. In this context, the simulation code acts as a large "black box" function f, transforming one input Dempster-Shafer structure on the line (also known as a random interval $\mathcal{A}$) into an output random interval $f(\mathcal{A})$. Our quantification of output uncertainty is then based on this output random interval. If some properties of f (perhaps monotonicity or other analytical properties) are known, then some information about $f(\mathcal{A})$ can be determined. But when f is a pure black box, we must resort to sampling approaches. In this paper, we present the basic formalism of a Monte Carlo approach to sampling a functionally propagated general random set, as opposed to a random interval. We show that the results of straightforward formal definitions are mathematically coherent, in the sense that bounding and convergence properties are achieved.

Keywords— Random sets, Dempster-Shafer theory, Monte Carlo sampling.

I. Introduction

We are interested in improving the risk and reliability analyses available for a certain class of technical systems.[1] These systems, such as nuclear facilities, are characterized by a high complexity such that our knowledge of their behavior is available primarily through large simulation codes. Moreover, many of their input parameters may be known only imprecisely.

Such imprecision lends itself to interval representations of parameter values, and thence to quantifying our uncertinay through Dempster-Shafer (random set) representations on the input space [20]. In this context, the simulation code acts as a large "black box" function, transforming an input random interval (Desmpter-Shafer structure on the line) into another. Our quantification of output uncertainty is then based on this output random interval.

If some properties of the black box function (perhaps monotonicity, global bounds, or some other analytical properties) are known, then some information about the output random interval can be determined [1], [15], [21], [22]. But when the function is a pure black box, we must resort to sampling approaches.

Monte Carlo sampling has been used extensively in Dempster-Shafer theory. But to our knowledge it has been limited to situations where the computational complexity of random set structure resulted in sampling approaches being required for estimating the statistical properties of individual random sets or combinations of random sets, for example their individual expectations [5], [16], their Dempster combinations [14], or more simply just their belief and plausibility measures [18].

In this paper we first outline the basic formalism of our approach to the use of Monte Carlo sampling to approximate the belief and plausibility measures of a random set propagated through a functional black box. We then show a bounding theorem which states that the resulting sample plausibility cannot overestimate the true output beliefs and plausibilities; and a convergence property stating that, given reasonable preconditions, both the sample plausibilty and belief converge to their true values in the limit of infinte samples.

II. Problem Description

A. Notation and Preliminaries

Let $X := \mathbb{R}^p, Y := \mathbb{R}^q$, and let $\mathcal{D}(X) := \{A\} \subseteq \mathcal{P}(X), \mathcal{D}(Y) := \{B\} \subseteq \mathcal{P}(Y)$, where $\mathcal{P}(X)$ is the power set of X, be σ-fields of nonempty, closed subsets of X and Y respectively. $\mathcal{D}(X)$ and $\mathcal{D}(Y)$ are thereby closed under complementation and countable union and intersection [2]. Generally, let $A \subseteq X, B \subseteq Y$ mean that $A \in \mathcal{D}(X), B \in \mathcal{D}(Y)$.

For two sets A_1, A_2, denote $A_1 \circ A_2 := A_1 \cap A_2 \neq \emptyset$, read as "$A_1$ intersects A_2"; and $A_1 \perp A_2 := A_1 \cap A_2 = \emptyset$, read as "$A_1$ and A_2 are disjoint".

Let $f: X \mapsto Y$ be a function measurable in $\mathcal{D}(Y)$ in that

$$\forall A \in \mathcal{D}(X), \quad f(A) \in \mathcal{D}(Y)$$

$$\forall B \in \mathcal{D}(Y), \quad f^{-1}(B) \in \mathcal{D}(X). \qquad (1)$$

Consider a class $\mathcal{C}_X := \{A_j\} \subseteq \mathcal{P}(X)$ on X, and define the induced class $\mathcal{C}_Y = f(\mathcal{C}_X) := \{f(A_j)\} = \{B_k\} \subseteq \mathcal{P}(Y)$ on Y, where $f(A) := \{f(x) : x \in A\}$.

Usually f takes a distinct input element $A_j \in \mathcal{C}_X$ to a distinct output element $f(A_j) \in \mathcal{C}_Y$. However, it may be the case that $\exists A_1, A_2 \in \mathcal{C}_X, f(A_1) = f(A_2) = B_k$. Thus generally $|\mathcal{C}_Y| \le |\mathcal{C}_X|$, and f partitions $\mathcal{C}_X$ via an equivalence relation denoted $\simeq$, where $A_1 \simeq A_2 := f(A_1) = f(A_2)$. Denote each equivalence class on $\mathcal{C}_X$ as $[B_k] := \{A_j \in \mathcal{C}_X : f(A_j) = B_k\}$, and denote the partition $\mathcal{C}_X / \simeq := \{[B_k]\}$.

Prepared for the 2002 Conference of the North American Fuzzy Information Processing Society (NAFIPS 02).

Cliff Joslyn: Distributed Knowledge Systems and Modeling Team, Modeling, Algorithms, and Informatics Group (CCS-3), Mail Stop B265, Los Alamos National Laboratory, Los Alamos, NM 87545, USA, joslyn@lanl.gov, http://www.c3.lanl.gov/~joslyn, (505) 667-9096.

Jon Helton: Arizona State University and Environment Decisions and WIPP Support, Sandia National Laboratories, Albquerque, NM USA, jchelto@sandia.gov, (505) 284-4808.

[1] http://www.sandia.gov/epistemic

The following lemma will be useful below.

Lemma 2 $\forall B \subseteq Y, \overline{f^{-1}\left(\overline{B}\right)} = f^{-1}(B)$, where $\overline{\{\cdot\}}$ is set complementation.

Proof: $\forall x \in X, x \in \{x' : f(x') \in B\} \leftrightarrow x \notin \{x' : f(x') \notin B\}$. Also, recall that $f^{-1}(B) = \{x : f(x) \in B\}$. Therefore

$$\begin{aligned}\overline{f^{-1}\left(\overline{B}\right)} &= \left\{x \notin f^{-1}\left(\overline{B}\right)\right\} = \{x \notin \{x' : f(x') \notin B\}\} \\ &= \{x : f(x) \in B\} = f^{-1}(B).\end{aligned}$$

■

We are working with observed data, which generally comes in the form of vectors or bags (collections which may have duplicates), and not sets (which may not). Thus we tend to use vector notation somewhat analogously to set notation, where possible hopefully without confusion. For example, for a vector $\vec{x} = \langle x_i \rangle, 1 \leq i \leq n$ of size $|\vec{x}| := n$, denote $x \in \vec{x} := \exists x_i, x = x_i$. For a vector $\vec{x}$, let $\mathbf{set}(\vec{x}) \subseteq X$ be the set obtained be eliminating duplicates from $\vec{x}$, so that $|\mathbf{set}(\vec{x})| \leq n$. And let $\vec{0} = \langle\rangle$ be that vector such that $n = 0$.

B. Random Sets and Dempster-Shafer Evidence Theory

We begin by laying out the basic notation of random set theory [13], [17], which is mathematically isomorphic to Dempster-Shafer evidence theory [6], [7].

Definition 3 (Finite Random Set) Assume an **evidence function (basic probability assignment)** $m_X : \mathcal{D}(X) \mapsto [0,1]$ where $m_X(\emptyset) = 0$ and $\sum_{A \subseteq X} m_X(A) = 1$. Then let $\mathcal{S}_X := \{\langle A_j, m_j \rangle : m_j > 0\}$ be a **finite random subset** of X generated by m_X, where $A_j \subseteq X, m_j := m_X(A_j)$, and $1 \leq j \leq N_X := |\mathcal{S}_X|$. Denote the **focal set** $\mathcal{F}_X := \{A_j : m_j > 0\}$, where each A_j is a **focal element**, and the **support** as

$$\mathrm{U}(\mathcal{S}) := \bigcup_{A_j \in \mathcal{F}_X} A_j. \tag{4}$$

◊

$\mathcal{S}_X$ has the **evidence measures** belief and plausibility $\mathrm{Bel}_X, \mathrm{Pl}_X : \mathcal{P}(\Omega) \mapsto [0,1]$, where $\forall A \subseteq X$

$$\mathrm{Bel}_X(A) = \sum_{A_j \subseteq A} m_j, \qquad \mathrm{Pl}_X(A) = \sum_{A_j \circ A} m_j. \tag{5}$$

We also have $\forall A \subseteq X$

$$\mathrm{Pl}_X(A) = 1 - \mathrm{Bel}_X\left(\overline{A}\right), \quad \mathrm{Bel}_X(A) = 1 - \mathrm{Pl}_X\left(\overline{A}\right), \tag{6}$$

$$\mathrm{Bel}_X(A) \leq \mathrm{Pl}_X(A). \tag{7}$$

Pl and Bel are fuzzy measures [23], and are thus monotone with

$$A \subseteq A' \rightarrow (\mathrm{Pl}_X(A) \leq \mathrm{Pl}_X(A'), \mathrm{Bel}_X(A) \leq \mathrm{Bel}_X(A')). \tag{8}$$

We also have the **Möbius transform**

$$\begin{aligned} m_X(A) &= \sum_{A' \subseteq A} (-1)^{|A-A'|} \mathrm{Bel}_X(A') \\ &= \sum_{A' \subseteq A} (-1)^{|A-A'|} \left(1 - \mathrm{Pl}_X\left(\overline{A'}\right)\right). \end{aligned} \tag{9}$$

In this way, each of m_X, Bel_X, and Pl_X is determined by any of the others.

Note that we can convolve the elements of the finite random set $\{\langle A_j, m_j \rangle\}$ to derive the structure $\langle \{A_j\}, \{m_j\} \rangle = \langle \mathcal{F}, m \rangle$, which is known as a **Dempster-Shafer body of evidence** [7]. In this way random set theory is isomorphic to Dempster-Shafer evidence theory [19].

C. Functionally Propagated Random Sets

We consider f as a "black box", and are intrested in how $\mathcal{S}_X$ is propagated by f to Y.

Definition 10 (Propagated Random Set) Let $\mathcal{S}_Y$ be the random set on Y induced by f and $\mathcal{S}_X$ with focal set $\mathcal{F}_Y = \{f(A_j)\} = \{B_k\} \subseteq \mathcal{D}(Y), 1 \leq k \leq N_Y := |\mathcal{F}_Y| \leq N_X$, evidence function $m_Y : \mathcal{D}(Y) \mapsto [0,1]$ such that

$$m_Y(B) := \begin{cases} \sum_{A_j : f(A_j) = B} m_X(A_j), & \exists A_j, f(A_j) = B \\ 0, & \text{otherwise} \end{cases}, \tag{11}$$

and plausibility and belief $\mathrm{Pl}_Y, \mathrm{Bel}_Y$. ◊

This definition has also been offered by Dubois and Prade [4], and is quite natural. First, it extends the case of standard random variables to random sets. Then, it reasonably accounts for all the input evidence in virtue of $\simeq$.

Proposition 12 $\forall B_k \in \mathcal{F}_Y$,

$$m_Y(B_k) = \sum_{A_j \in [B_k]} m_X(A_j).$$

Proof: Trivial from (11). ■

The following conservation property is both useful and comforting.

Theorem 13 $\forall B \subseteq Y$,

$$\mathrm{Pl}_Y(B) = \mathrm{Pl}_X\left(f^{-1}(B)\right), \quad \mathrm{Bel}_Y(B) = \mathrm{Bel}_X\left(f^{-1}(B)\right).$$

Proof: Consider a set $B \subseteq Y$. For any $B_k \in \mathcal{F}_Y$, if $B_k \circ B$, then $\forall A_j \in [B_k], A_j \circ f^{-1}(B)$. Therefore

$$\forall A_j \in \left(\bigcup_{B_k \circ B} [B_k]\right), \quad A_j \circ f^{-1}(B).$$

Furthermore, there can be no other $A_0 \notin \left(\bigcup_{B_k \circ B} [B_k]\right)$ with $A_0 \circ f^{-1}(B)$, since otherwise there would be another $B_0 = f(A_0)$ with $B_0 \circ B$. Therefore, in virtue of (12) and the fact that $[B_k] \in \mathcal{F}_X / \simeq$ is an equivalence class, we have

$$\begin{aligned} \mathrm{Pl}_X\left(f^{-1}(B)\right) &= \sum_{A_j \circ f^{-1}(B)} m_X(A_j) \\ &= \sum_{A_j \in \left(\bigcup_{B_k \circ B} [B_k]\right)} m_X(A_j) \\ &= \sum_{B_k \circ B} \sum_{A_j \in [B_k]} m_X(A_j) \\ &= \sum_{B_k \circ B} m_Y(B_k) = \mathrm{Pl}_Y(B). \end{aligned}$$

The belief result follows from (2) and (6):

$$\begin{aligned}\mathrm{Bel}_Y(B) &= 1-\mathrm{Pl}_Y\left(\overline{B}\right) = 1-\mathrm{Pl}_X\left(f^{-1}\left(\overline{B}\right)\right)\\ &= 1-\left(1-\mathrm{Bel}_X\left(\overline{f^{-1}\left(\overline{B}\right)}\right)\right)\\ &= \mathrm{Bel}_X\left(\overline{f^{-1}\left(\overline{B}\right)}\right) = \mathrm{Bel}_X(f^{-1}(B)).\end{aligned}$$

■

The result is illustrated in Fig. 1, showing all the relationships among the various components.

III. Sampling Approach

We wish to characterize Pl_Y and Bel_Y in terms of $\mathcal{S}_X$ and f. However, while we assume that we know $\mathcal{S}_X$ completely, we will not know about the general structure of f, but will only know some sample evaluations.

A. Observations in Sample

Assume that input points $x_i \in X, 1 \le i \le n$, are selected without replacement, and the function f evaluated to determine each $y_i := f(x_i)$. Then let each $s_i := \langle x_i, y_i\rangle$ be an **observation**, and $S := \{s_i\} = \{\langle x_i, y_i\rangle\}$ the **observation record**.

Let $x \in S$ denote that x has been observed in S: $x \in S := \exists \langle x_i, y_i\rangle \in S, x = x_i$, and similarly for $y \in S$. Since f is generally many to one, there may be points observed in the output which are mapped to by points in the input which are themselves not observed: $\exists y_i \in S, \exists x \in X, f(x) = y_i$ and $x \notin S$. Furthermore, because the x_i are selected without replacement, the collection of the x_i form a set of n distinct points. However, the collection of output points y_i may contain duplicates, and is thus generally not a set, but rather a bag or vector.

In this way sampled points necessarily may lose crucial information concerning the inverse relation f^{-1}. But nevertheless they are all we have. So we must always clearly distinguish pairs $\langle x, f(x)\rangle$ which are *functionally related* in f from pairs $\langle x_i, f(x_i)\rangle \in S$ which have been *actually observed* "in sample".

Definition 14 (Observations in Sample) Denote

$$S(A) := \{x_i \in S : x_i \in A\}$$

as the set of x_i values within a subset $A \subseteq X$ which have been observed in S. However, for a subset $B \subseteq Y$, the collection of observed y_i output values is a vector, denoted $\vec{S}(B) := \langle y_i \in S : y_i \in B\rangle$. Define

$$\widehat{f}^{-1}\left(\vec{S}(B)\right) := \left\{x_i : y_i \in \vec{S}(B)\right\}$$

as the **inverse in sample of** B: the set of input values actually observed in correspondence with the output values $\vec{S}(B)$ in the output region B. ◇

Note that $\left|\widehat{f}^{-1}\left(\vec{S}(B)\right)\right| \le n$. Moreover, for a given output region B, fewer input points may be present in its inverse in sample than are contained in the actual functional inverse of the output points seen in B. And of course, these are always fewer than those in the actual functional inverse of the region B as a whole.

Proposition 15

$$\widehat{f}^{-1}\left(\vec{S}(B)\right) \subseteq f^{-1}\left(\mathbf{set}\left(\vec{S}(B)\right)\right) \subseteq f^{-1}(B) \subseteq X.$$

Proof: The first relation follows from f being many to one. The second follows from the fact that $\vec{S}(B) \subseteq B$.

■

An example is shown in Fig. 2 for $\mathcal{F}_X = \{A_1, A_2\}, A_1 \supseteq \{x_1, x_2\}, A_2 \supseteq \{x_3, x_4\}$ and $f(x_2) = f(x_3)$. Assume an observation record $S = \{\langle x_1, y_1 = f(x_1)\rangle, \langle x_2, y_2 = f(x_2)\rangle\}$. Note that $\langle x_3, f(x_3)\rangle$ has *not* been observed. Then for B as shown, we have $\vec{S}(B) = \langle y_1, y_2\rangle$, and

$$\begin{aligned}\widehat{f}^{-1}\left(\vec{S}(B)\right) &= \{x_1, x_2\}\\ &\subseteq \left[f^{-1}\left(\mathbf{set}\left(\vec{S}(B)\right)\right) \supseteq \{x_1, x_2, x_3\}\right]\\ &\subseteq \left[f^{-1}(B) \supseteq \{x_1, x_2, x_3, x_4\}\right] \subseteq X.\end{aligned}$$

B. Sample Plausibility and Belief

Since we do not know f, therefore we cannot determine the output plausibility or belief directly. But we can propose a "sample plausibility" and belief on Y as an estimator.

The sample plausibility is simply the input plausibility of those points which are seen in a sample.

Definition 16 (Sample Plausibility) Given $\mathcal{S}_X$ and S, let $\widehat{\mathrm{Pl}}_Y : \mathcal{D}(Y) \mapsto [0,1]$, with

$$\widehat{\mathrm{Pl}}_Y(B) := \mathrm{Pl}_X\left(\widehat{f}^{-1}\left(\vec{S}(B)\right)\right), \quad B \subseteq Y.$$

◇

In words, for a given region $B \subseteq Y$ of the output space, its sample plausibility $\widehat{\mathrm{Pl}}_Y(B)$ is the plausibility in the input random set of the collection of the x values observed in association with whichever y were actually observed in B.

It is tempting to define the sample belief similarly as $\mathrm{Bel}_X\left(\widehat{f}^{-1}\left(\vec{S}(B)\right)\right)$, but this is troublesome: since $\widehat{f}^{-1}\left(\vec{S}(B)\right)$ is a finite collection of points in X, therefore there can be no $A_j \subseteq \widehat{f}^{-1}\left(\vec{S}(B)\right)$, and thus $\forall B \subseteq X, \mathrm{Bel}_X\left(\widehat{f}^{-1}\left(\vec{S}(B)\right)\right) \equiv 0$.

Instead, introduce the following from (6):

Definition 17 (Sample Belief) Given $\mathcal{S}_X$ and S, let $\widehat{\mathrm{Bel}}_Y : \mathcal{D}(Y) \mapsto [0,1]$, with

$$\widehat{\mathrm{Bel}}_Y(B) := 1 - \widehat{\mathrm{Pl}}_Y\left(\overline{B}\right).$$

◇

It must be emphasized that $\widehat{\mathrm{Pl}}_Y$ and $\widehat{\mathrm{Bel}}_Y$ are actually not themselves plausibility and belief measures, but rather estimators or approximations of them.

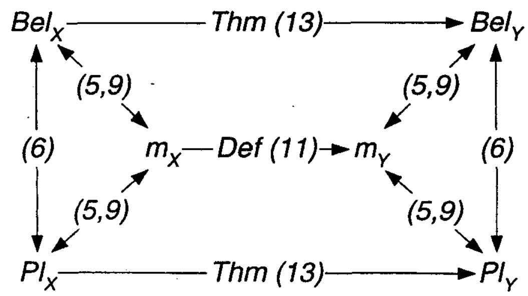

Fig. 1. Relationships among the components of (13).

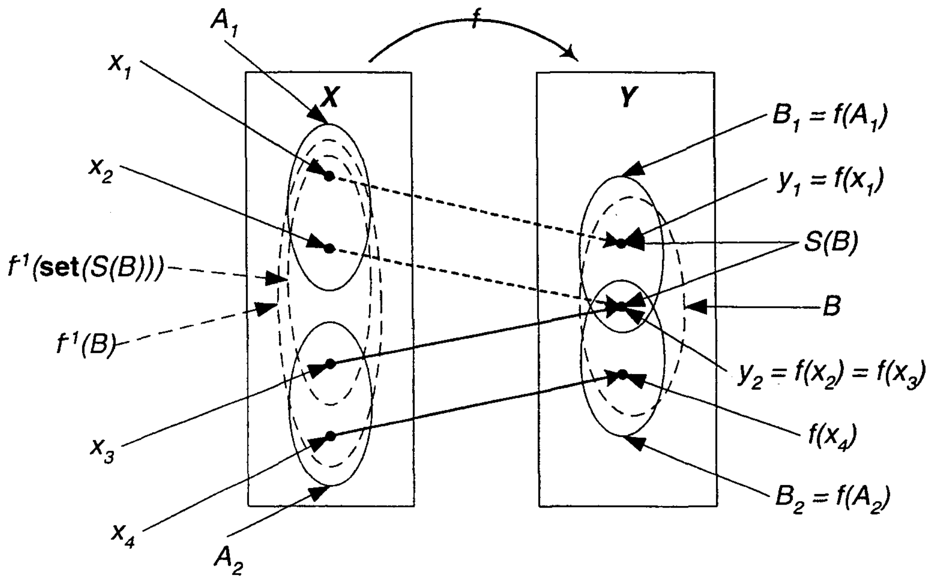

Fig. 2. Example of observations in sample.

Theorem 18 $\widehat{\mathrm{Pl}}_Y$ is not a plausibility measure and $\widehat{\mathrm{Bel}}_Y$ is not a belief measure.

Proof: A simple, but extreme, counterexample will suffice. Consider a $B \subseteq Y$ where $B \circ B_k$ for some B_k, but $\vec{S}(B) = \vec{\emptyset}$; in other words, there just aren't any sample points hitting B. So $\widehat{\mathrm{Pl}}_Y(B) = 0$, although $\mathrm{Pl}_Y(B) > 0$. Now since $\vec{S}(B) = \vec{\emptyset}$, therefore $\vec{S}\left(\overline{B}\right) = \langle y_1, y_2, \ldots, y_n \rangle$, that is, the complete sample record. Thus $\widehat{f}^{-1}\left(\vec{S}(\overline{B})\right) = \mathrm{set}\left(\langle x_1, x_2, \ldots, x_n \rangle\right) \neq \emptyset$, and $\widehat{\mathrm{Bel}}_Y(B) = 1 - \widehat{\mathrm{Pl}}_Y\left(\overline{B}\right) \geq 0$. So $\widehat{\mathrm{Pl}}_Y(B) = 0 \leq \widehat{\mathrm{Bel}}_Y(B)$, violating (7) . Moreover, $\widehat{\mathrm{Pl}}_Y(B) = \widehat{\mathrm{Bel}}_Y(B)$ only if $\widehat{\mathrm{Pl}}_Y\left(\overline{B}\right) = 1$, which only occurs if $\forall A_j \in \mathcal{F}_Y, \exists x_i \in S, x_i \in A_j$, which very well might not be the case. Therefore $\widehat{\mathrm{Pl}}_Y$ and $\widehat{\mathrm{Bel}}_Y$ cannot be plausibility and belief measures. ∎

IV. Bounding and Convergence

We first show a weak bounding result on sample plausibility, and then propose a stronger convergence result in the limit of infinite samples.

A. Bounding

It is intuitive that any sampling approach would tend to underestimate our uncertainty measure: since there will be points not observed, there may be focal elements not seen in the output, and thus no opportunity to add their masses into our estimator. And this does, indeed, prove to be the case for the sample plausibility.

Theorem 19 $\widehat{\mathrm{Pl}}_Y(B) \leq \mathrm{Pl}_Y(B)$.

Proof: From (13) and (16), we need to show that

$$\mathrm{Pl}_X\left(\widehat{f}^{-1}\left(\vec{S}(B)\right)\right) \leq \mathrm{Pl}_X\left(f^{-1}(B)\right).$$

The result then follows from (8) and (15). ∎

TABLE I

Example evidence measure values.

B	$\text{Pl}_Y(B)$	$\widehat{\text{Pl}}_Y(B)$	$\text{Bel}_Y(B)$	$\widehat{\text{Bel}}_Y(B)$
B'	1.0	0.5	0.8	0.2
B''	0.3	0.3	0.0	0.5

However, in general there is no relation between $\widehat{\text{Bel}}_Y(B)$ and either $\text{Bel}_Y(B)$ or even $\widehat{\text{Pl}}_Y(B)$. Consider the example in Fig. 3, showing the input random set

$$\mathcal{S} = \{\langle A_1, 0.3\rangle, \langle A_2, 0.5\rangle, \langle A_3, 0.2\rangle\}.$$

As in Fig. 2, a number of individual points in X are shown, and their functional images in Y. In this case, only x_2, x_3, and x_4 have been sampled. Also, $f(x_1) = f(x_2)$, so that $B_2 \subseteq B_1$, whereas $A_2 \perp A_1$. The resulting evidence measures are shown in Tab. I for the two output sets B' and B'' shown. We can see that

$$\widehat{\text{Bel}}_Y(B') < \text{Bel}_Y(B'), \quad \widehat{\text{Bel}}_Y(B'') > \text{Bel}_Y(B''),$$

$$\widehat{\text{Bel}}_Y(B') < \widehat{\text{Pl}}_Y(B'), \quad \widehat{\text{Bel}}_Y(B'') > \widehat{\text{Pl}}_Y(B'').$$

This is a reflection of the sense in which plausibility is the "primary" measure in the $\langle \text{Bel}, \text{Pl}\rangle$ pair, similar to the role of the possibility measure Π in possibility theory, given the existence of the possibility/necessity pair $\langle \eta, \Pi\rangle$.

B. Convergence

Despite the weak bounding result, there is a stronger convergence property: given reasonable assumptions, in the limit of an infinite number of sample points, both the sample plausibility and belief approach their true values. We present the following propostion without proof, which is under development for a future publication [12].

Proposition 20 Assume an input random set $\mathcal{S}$, a function $f: X \mapsto Y$ which is measurable in the sense of (1) and appropriately continuous, and an observation record S with $|S| = n$. Furthermore, assume that sampling of input regions is such that $\forall A \in \mathcal{D}(X)$, in the limit of infinite samples $A - S(A)$ is of measure zero. Then $\forall B \in \mathcal{D}(Y)$,

$$\lim_{n\longrightarrow\infty} \widehat{\text{Pl}}_Y(B) = \text{Pl}_Y(B), \quad \lim_{n\longrightarrow\infty} \widehat{\text{Bel}}_Y(B) = \text{Bel}_Y(B).$$

V. Conclusions and Further Work

We have demonstrated that Monte Carlo sampling of functionally propagated random sets is mathematically coherent, in the sense that the obvious definitions lead to appropriate bounds and convergence. However, this is not to say that it is necessarily a *useful* method for the problems we are interested in.

We are currently also working on the following issues, which will be detailed in future papers:

- In our real problems, we work exclusively not with general random sets on X, but with functions mapping statistical collections of intervals and probability distributions on a moderately high dimensional space [20]. Thus we need to translate these results into the more specialized cases first of general random intervals $\mathcal{A}$ [3], [4], [8], [9], [10], [11], where $X, Y = \mathbb{R}$, and the $\mathcal{D}(X), \mathcal{D}(Y)$ are Borel sets; then to multi-dimensional random hyper-intervals where $X = \mathbb{R}^p$ again, but $\mathcal{D}(X)$ are p-dimensional hyperrectangles; and finally to the hybrid case involving both random intervals and probability distributions.
- While the convergence result is important, of critical concern is the *rate* of convergence. And here we have cause for concern, especially given that $|\mathcal{F}_X|$ grows exponentially with p. We are thus actively considering a number of methods to accelerate convergence. Some of these are generalizations to the random set case of traditional acceleration approaches, including stratified, importance, and sensitivity sampling and approaches based on optimization of f. But others are unique to the random set approach, as they reflect the *structural* nature of $\mathcal{F}_X$.
- Related both to the convergence rate and the amount of error present in any (probably necessarily severe) undersampling, is the fact that there are two different components to the underestimation of $\text{Pl}_Y(B)$:

1. That resulting from undersampling of a particular input focal element A_j.
2. That resulting from f being many-to-one, so that input focal elements present in the inverse of B are not touched by the inverse in sample: in other words, there may be sample pairs $\langle x_i, y_i\rangle$ such that there exists other $x' \neq x_i$ such that $f(x') = y_i$, but $\langle x', y_i\rangle$ is not in the sample record S, as discussed in Sec. III-A.

Note that while 1 is present in traditional Monte Carlo sampling, 2 is unique to the random set approach.

- Since our problems are mostly in the risk analysis domain, a true bounding result where some estimator $\widetilde{\text{Pl}}_Y(B)$ was available with $\widetilde{\text{Pl}}_Y(B) \geq \text{Pl}_Y(B)$ would be truly useful. If such a measure were moreover computationally tractible and quickly convergent to $\text{Pl}_Y(B)$, then the general approach of propagating random intervals for risk analysis could be considerably more feasible.
- Finally, we have developed substantial experience with numerical implementations of the random hyper-interval case both for test problems and some simple real problems. We will be discussing these results in print, as they reveal empirical insights into the nature of these processes.

Acknowledgements

This work is supported by a research grant from Sandia National Laboratories as part of the Department of Energy Accelerated Strategic Computing Initiative (ASCI).

We would also like to thank all the other members of Sandia's Epistemic Uncertainty project[2] for their intense interaction, discussion, review, and guidance of this work: Bill Oberkampf, Scott Ferson, Martin Pilch, Kari Sentz, and Steve Wojtkiewicz.

Finally, we would like to thank Paul Pedersen of the Los Alamos National Laboratory, for giving us the encourage-

[2] `http://www.sandia.gov/epistemic`

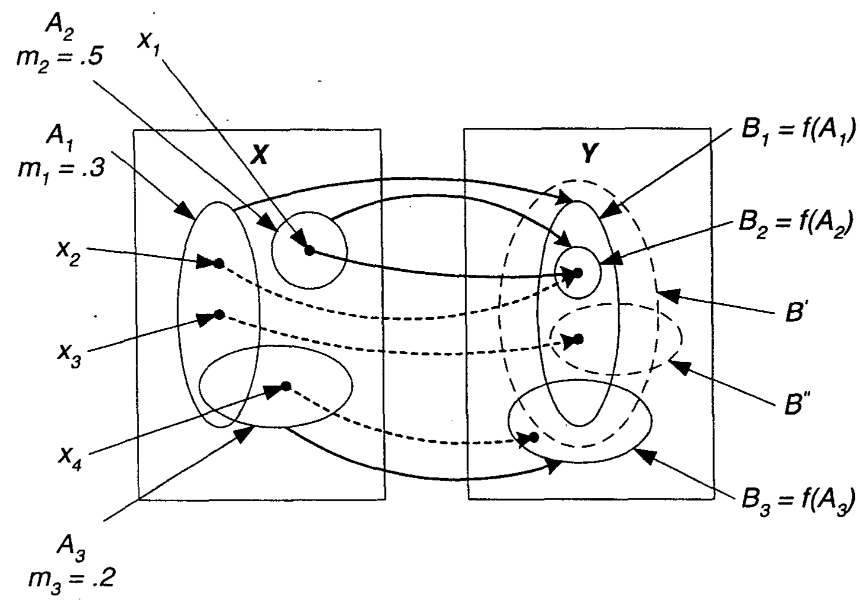

Fig. 3. Example of a propagated random set.

ment to muscle through the proof of (19).

References

[1] Alt, R and Lamotte, JL: (2001) "Experiments on the Evaluation of Functional Ranges Using a Random Interval Arithmetic", *Mathematics and Computers in Simulation*, v. **56**, pp. 17-34

[2] Capiński, Marek and Kopp, Ekkehard: (1999) *Measure, Integral and Probability*, Springer-Verlag, Berlin

[3] Dempster, AP: (1968) "Upper and Lower Probabilities Generated by a Random Interval", *Annals of Mathematical Statistics*, v. **39**:3, pp. 957-966

[4] Dubois, Didier and Prade, Henri: (1991) "Random Sets and Fuzzy Interval Analysis", *Fuzzy Sets and Systems*, v. **42**, pp. 87-101

[5] García, D; Lubiano, MA; and Alonso, MC: (2001) "Estimating the Expected Value of Fuzzy Random Variables in the Stratified Random Sampling from Finite Populations", *Information Sciences*, v. **138**, pp. 165-184

[6] Goutsias, J; Mahler, Ronald PS; and Nguyen, HT, eds.: (1998) *Random Sets: Theory and Applications*, Springer-Verlag

[7] Guan, J and Bell, DA: (1992) *Evidence Theory and its Applications*, v. **1-2**, North Holland

[8] Joslyn, Cliff: (1997) "Measurement of Possibilistic Histograms from Interval Data", *Int. J. General Systems*, v. **26**:1-2, pp. 9-33

[9] Joslyn, Cliff: (1997) "Possibilistic Normalization of Inconsistent Random Intervals", *Advances in Systems Science and Applications*, v. **special**, ed. Wansheng Tang, pp. 44-51, San Marcos, TX

[10] Joslyn, Cliff: (1997) "Distributional Representations of Random Interval Measurements", in: *Uncertainty Analysis in Engineering and the Sciences*, ed. Bilal Ayyub and Madan Gupta, pp. 37-52, Kluwer

[11] Joslyn, Cliff: (2001) "Measures of Distortion in Possibilistic Approximations of Consistent Random Sets and Intervals", in: *Proc. 2001 Joint Conf. North American Fuzzy Information Processing Society and the International Fuzzy Systems Association*, pp. 1735-1740

[12] Joslyn, Cliff; Helton, John C; and Pedersen, Paul: (2002) "Monte Carlo Estimation of Propagated Random Sets", in preparation for *Int. J. Approximate Reasoning*

[13] Kendall, DG: (1974) "Foundations of a Theory of Random Sets", in: *Stochastic Geometry*, ed. EF Harding and DG Kendall, pp. 322-376, Wiley, New York

[14] Kreinovich, V; Bernat, Andrew; and Borrett, W et al.: (1994) "Monte-Carlo Methods Make Dempster-Shafer Formalism Feasible", in: *Advances in the Dempster-Shafer Theory of Evidence*, ed. R Yager *et al.*, pp. 175-191

[15] Kreinovich, V and Trejo, Raúl: (2001) "Error Estimations for Indirect Measurements: Randomized vs. Deterministic Algorithms for "Black Box" Programs", in: *Handbook on Randomized Computing*, ed. S. Rajasekaran et al., pp. 673-729, Kluwer

[16] Lubiano, MA and Gil, MA: (1999) "Estimating the Expected Value of Fuzzy Random Variable in Random Samplings from Finite Populations", *Statistical Papers*, v. **40**, pp. 277-295

[17] Matheron, G: (1975) *Random Sets and Integral Geometry*, Wiley, New York

[18] Moral, S and Wilson, Nic: (1996) "Importance Sampling Monte-Carlo Algorithms for the Calculation of Dempster-Shafer Belief", in: *Proc. IPMU-96*, v. **3**, pp. 1337-1344

[19] Nguyen, HT and Wang, Tonghui: (1998) "Belief Functions and Random Sets", in: *Random Sets: Theory and Applications*, pp. 243-256, Springer-Verlag

[20] Oberkampf, WL; Helton, JC; and Sentz, Kari: (2001) "Mathematical Representation of Uncertainty", in: *Proc. Non-Deterministic Approaches Forum*

[21] Tonon, Fulvio; Bernardini, Alberto; and Mammino, A: (2000) "Reliability Analysis of Rock Mass Response by Means of Random Set Theory", *Reliability Engineering and System Safety*, v. **70**, pp. 263-282

[22] Tonon, Fulvio; Bernardini, Alberto; and Mammino, A: (2000) "Determination of Paramaters Range in Rock Engineering by Means of Random Set Theory", *Reliability Engineering and System Safety*, v. **70**, pp. 241-261

[23] Wang, Zhenyuan and Klir, George J: (1992) *Fuzzy Measure Theory*, Plenum Press, New York