Joslyn, Cliff A: (2001) "Measures of Distortion in Possibilistic Approximations of Consistent Random Sets and Intervals", in: Proc. 2001 Joint International Fuzzy Systems Association World Congress and North American Fuzzy Information Processing Society Int. Conf., pp. 1735-1740, https://doi.org/10.1109/NAFIPS.2001.943814

# Measures of Distortion in Possibilistic Approximations of Consistent Random Sets and Intervals

Dr. Cliff Joslyn Los Alamos National Laboratory Los Alamos, New Mexico, USA joslyn@lanl.gov www.c3.lanl.gov/~joslyn

## Abstract

The available of probability or possibility measures for random set (Dempster-Shafer evidence theoretical) structures are highly desirable. Probabilistic conditions involve disjointness or specificity, while possibilistic conditions involve consonance of the underlying focal elements. Consistency results in possibilistic distributions, but not measures, but then a unique approximation is available. Especially in random interval measurement situations, this condition is common. In this paper we develop some of the mathematical ideas necessary to develop a measure of the distortion introduced by this consonant approximation of a consistent random set.

#### 1. Introduction

Random set measurement (mathematically equivalent to the derivation of Dempster-Shafer bodies of evidence from empirical sources) is a welljustified method for the empirical determinination of, on the one hand, specialized point-trace forms like possibility distributions and other normalized fuzzy sets like fuzzy numbers and intervals; and, on the other hand, specialized evidence measures (belief and plausibility measures) like possibility measures [5, 7].

An issue in the use of this method is the underlying class structure of the focal set of observed subsets. This structure can be completely general, or have a special structure such as specific (singleton subsets), disjoint, rings (intersecting subsets), consistent (global intersecting subsets), or consonant (completely nested subsets) [4]. Depending on the structure present, the resulting point traces and evidence measures have different properties. When consonant, the plausibility measure is naturally a possibility measure. In the more general case of consistency, while the plausibility measure is not a natural possibility measure, the trace is a natural possibility distribution, and a unique and welljustified possibilistic approximation of the plausibility measure is available. Finally, consistent approximations of inconsistent random sets are also available [6].

Consonant random sets provide a good empirical basis for applications of possibility theory. And yet, consistent but non-consonant random sets are a much more common case, for example in problems requiring the combination of interval measurements [3, 8]. Possibilistic approximations of these are not only well justified, they also bring great computational advantages, since they are representable in linear, rather than exponential, computational space and time. But of course such advantages are bought at the price of approximation: the "less consonant" the underlying consistent random set, the more the resulting approximated possibility measure distorts the empirical plausibility measure actually generated by the measured evidence.

In this paper we develop some of the mathematical ideas necessary to develop a measure of the distortion introduced by a consonant approximation of a consistent random set.

#### 2. Random Sets and Fuzzy Measures

Given a finite universe  $\Omega := \{\omega_i\}, 1 \leq i \leq n$  and evidence function (basic probability assignment)  $m: 2^{\Omega} \mapsto [0, 1]$ , let  $\mathcal{S} := \{\langle A_j, m_j \rangle : m_j > 0\}$  be a finite random set generated by m, where  $\langle \cdot \rangle$  is a vector,  $A_j \subseteq \Omega, m_j := m(A_j)$ , and  $1 \leq j \leq N :=$  $|\mathcal{S}| \leq 2^n - 1$ . Let  $\vec{m} := \langle m_j \rangle$  be an N-vector. Denote the focal set as  $\mathcal{F}(\mathcal{S}) := \{A_j : m_j > 0\}$ , where each  $A_j$  is a focal element, and the core as

$$\mathbf{C}(\mathcal{S}) = \mathbf{C}(\mathcal{F}(\mathcal{S})) := \bigcap_{A_j \in \mathcal{F}(\mathcal{S})} A_j.$$

 $\mathcal S$  has belief and plausibility measures Bel, Pl:  $2^\Omega\mapsto [0,1],$  where  $\forall A\subseteq\Omega$ 

$$\operatorname{Bel}(A) = \sum_{A_j \subseteq A} m_j, \qquad \operatorname{Pl}(A) = \sum_{A_j \not \downarrow A} m_j \quad (1)$$

and  $A \perp B := A \cap B = \emptyset$ . Given two random sets S, S' on  $\Omega$ , then random set inclusion is defined as [1]  $S \subseteq S' := \forall A \subseteq \Omega, \operatorname{Pl}(A) \leq \operatorname{Pl}'(A)$ .

Pl and Bel have the following linear algebraic formulations:

$$\vec{\mathrm{Bel}} = \mathbf{G}^{\subseteq} \vec{m}^T, \quad \vec{\mathrm{Pl}} = \mathbf{G}^{\cap} \vec{m}^T,$$

where:

- Bel :=  $\langle Bel(A) \rangle$  and  $Pl := \langle Pl(A) \rangle$  are  $2^n 1$ -vectors of the belief and plausibility measure values  $\forall A \subseteq \Omega$ ;
- $\mathbf{G}^{\subseteq}$  and  $\mathbf{G}^{\cap}$  are  $2^n 1 \times 2^n 1$  matrices where for  $1 \leq j,k \leq 2^n 1$

$$\mathbf{G}_{j,k}^{\subseteq} = \begin{cases} 1, & A_j \subseteq A_k \\ 0, & \text{otherwise} \end{cases}$$
$$\mathbf{G}_{j,k}^{\cap} = \begin{cases} 1, & A_j \not\perp A_k \\ 0, & \text{otherwise} \end{cases}$$

Note that  $\mathbf{G}^{\subseteq}$  can be derived as the tensor product

$$\mathbf{G}^{\subseteq} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{\otimes n},$$

and  $\mathbf{G}^{\cap} = 1 - \mathbf{G}^{\prime \subseteq}$ , where  $\mathbf{G}^{\prime \subseteq}$  is  $\mathbf{G}^{\subseteq}$  reflected through the columns.

Bel and Pl are fuzzy measures  $\nu: 2^{\Omega} \mapsto [0, 1]$ , and are thereby monotonic with  $A \subseteq B \to \nu(A) \leq \nu(B)$ , and normal with  $\nu(\Omega) = 1$ . A fuzzy measure is decomposable in a conorm  $\sqcup: [0, 1]^2 \mapsto [0, 1]$  (a commutative, monotonic, associative operator with identity 0) if

$$\nu(A \cup B) \sqcup \nu(A \cap B) = \nu(A) \sqcup \nu(B).$$

For any fuzzy measure  $\nu$ , define the trace  $\rho_{\nu}: \Omega \mapsto [0, 1]$  as the singleton measure value  $\rho_{\nu}(\omega_i) := \nu(\{\omega_i\})$ . For a random set S, denote its trace as  $\rho^S := \rho_{\text{Pl}}$ , so that

$$\rho^{\mathcal{S}}(\omega_i) = \sum_{A_j \ni \omega_i} m_j. \tag{2}$$

Denote the *n*-vector  $\vec{\rho} = \langle \rho_i^S \rangle := \langle \rho^S(\omega_i) \rangle$ . Decomposability is equivalent to distributionality:

**Theorem 1** [4] A fuzzy measure  $\nu$  is decomposable in  $\sqcup$  if and only if  $\forall A \subseteq \Omega, \nu(A) = \bigcup_{\omega_i \in A} \rho_{\nu}(\omega_i)$ 

where we use operator notation for  $\sqcup$  in virtue of its associativity. Note that normalization of  $\nu$  results in normalization of the trace  $\bigsqcup_{\omega_i \in \Omega} \rho(\omega_i) = 1$ .

Any fuzzy measure determines its distribution:  $\nu \rightarrow \rho_{\nu}$ . Decomposable fuzzy measures are moreover uniquely determined by their distributions, so that  $\nu \leftrightarrow \rho_{\nu}$ . This is a great advantage computationally, since  $|\nu|$  grows with  $2^n$ , but  $|\rho_{\nu}|$  grows only with *n*. Thus a distributional fuzzy measure can be characterized by vastly less information than its non-distributional cousins.

For a random set S, define the numerical aggregator as the  $n \times N$  matrix  $\mathbf{H} := [h_{ij}]$ , where  $h_{ij} := \chi_{A_j}(\omega_i)$  and  $\chi_{A_j}$  is the characteristic function of  $A_j$ , so that

$$\chi_{A_j}(\omega_i) := \left\{ egin{array}{ccc} 1, & \omega_i \in A_j \ 0, & \omega_i 
otin A_j \end{array} 
ight., A_j \in \mathcal{F}(\mathcal{S}), \quad \omega_i \in \Omega \end{array} 
ight.$$

The j'th column of **H** is the vector representation of the characteristic function of  $A_j$ , so that **H** is a boolean matrix with no column repeated. Letting  $\mathbf{H}(i)$  be the *i*'th row of **H**, an *N*-vector, then

$$\rho_i = \mathbf{H}(i) \cdot \vec{m}^T, \qquad \vec{\rho}^T = \mathbf{H}\vec{m}^T \tag{3}$$

so that the left and right sides of (3) are the vector and matrix representations of (2) and (1) respectively.

We now introduce some example structures to be used below. Let  $\Omega = \{x, y, z\}$  so that n = 3. The structure of  $2^{\Omega}$  is shown in a three dimensional boolean hypercube in Fig. 1, with the subsets of  $\Omega$  related by inclusion relations. The example random set  $S = \{\langle \{x, z\}, .3 \rangle, \langle \{y, z\}, .2 \rangle, \langle \{x\}, .5 \rangle\}$  is shown on the left of Fig. 2, with the Bel and Pl measures on the center and right, where zero values are not shown. For example, we have  $Bel(\{z\}) =$  $0, Pl(\{x, z\}) = .5 = \rho(z) = .5$ .

## 3. Probability and Possibility

When  $\forall A_j \in \mathcal{F}(S), |A_j| = 1$  then S is called specific, and  $\Pr := \operatorname{Bel} = \operatorname{Pl}$  is a probability measure. Pr is decomposable for  $\sqcup = +_b$ , where  $x +_b y := (x + y) \wedge 1$  and  $\wedge$  is the minimum operator. Normalization of Pr implies that  $+_b = +$  here, so that

$$Pr(A \cup B) + Pr(A \cap B) = Pr(A) + Pr(B)$$
$$Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$$



Figure 2. (Left) An example random set. (Center) Bel. (Right) Pl and  $\rho$ .



Figure 1. The structure of  $2^{\Omega}$ .

and  $p := \rho$  becomes a probability distribution with  $\Pr(A) = \sum_{\omega_i \in A} p(\omega_i).$ 

When, for an arbitrary permutation of the j, and letting  $A_0 = \emptyset$ ,  $\forall A_j, A_{j-1} \subseteq A_j$ , then S is called consonant and  $\Pi := \operatorname{Pl}$  is a possibility measure.  $\Pi$  is decomposable for the maximum operator  $\lor$ , so that  $\Pi(A \cup B) \lor \Pi(A \cap B) = \Pi(A) \lor \operatorname{Pr}(B)$ , and from the monotonicity of  $\nu$  and the fact that  $A \cap B \subseteq$  $A \cup B, \Pi(A \cup B) = \Pi(A) \lor \Pi(B). \ \pi := \rho$  is now a possibility distribution with  $\Pi(A) = \bigvee_{\omega_i \in A} \pi(\omega_i)$ . Further, we have

$$m(A_i) = \pi_i - \pi_{i+1},$$
 (4)

where  $A_i$  is that  $A_j$  such that  $A_j = \{\omega_1, \omega_2, \dots, \omega_i\}$ , and  $\pi_{n+1} := 0$  by convention.

## 4. Consistent Random Sets and Their Consonant Approximations

Other fuzzy measures available on random sets are decomposable, but it is strongly suggested [4] that possibility and probability are the only measures for which not only are the *measure* values Pl(A) determined by the singletion measure values  $\rho_i = Pl(\{\omega_i\})$ , but so also is the identity of the *focal* elements  $A_j$  from the universe elements  $\omega_i$ .

Therefore, the availability of probabilistic or possibilistic representations of random sets is a very valuable thing to do. However, the constraints of specificity or consonance are severe, rarely occurring in a pure form in generalized measurement situations. Thus we wish to identify where reasonable approximations which produce probability and possibility measures might be available, and at what cost in terms of accuracy and computation.

When  $C(S) \neq \emptyset$ , then S is called consistent. Of course, all consonant random sets are consistent, with  $C(S) = \text{that } A_j$  such that  $|A_j|$  is minimal. Consistent random sets have the following useful property:

**Theorem 2** [2] A random set S is consistent if and only if  $\bigvee_{\omega_i \in \Omega} \rho^S(\omega_i) = 1$ .

So given a consistent random set S, its plausibility measure Pl is not a possibility measure, but its trace  $\rho^S$  is a possibility distribution, which we identify as  $\pi^* := \rho^S$ . Since, in turn, every possibility distribution determines a unique possibility measure, it is possible to construct a new possibility measure  $\Pi^*$ based on  $\rho^S$  from  $\forall A \subseteq \Omega, \Pi^*(A) := \bigvee_{\omega_i \in A} \pi^*(\omega_i)$ , and from (4) a new evidence function  $m^*$  and consonant random set  $S^*$ . Dubois and Prade have shown:

**Theorem 3** [1] If S is a consistent random set, then  $S^* \subseteq S$ .

While we are considering general finite random sets, we have a special interest in random intervals defined as statistical collections of half-open intervals of  $\mathbb{R}$  [5]. These structures are common in uncertain measurement problems. Note that while random intervals are defined on a continuous space, they are finite collections of intervals, and thus treatable as finite random sets. An example is shown in Fig. 3, including the resulting possibilistic histogram.



Figure 3. (Bottom) A consistent random interval. (Top) Its possibilistic histogram.

We now wish to characterize finite consistent random sets in general. For simplicity, consider a general consistent random set S which is complete in the sense that there is a unique singleton focus, and that there is positive evidentiary value for all subsets containing the focus:

$$\exists ! \omega^* \in \Omega, \mathbf{C}(\mathcal{S}) = \{\omega^*\},$$
$$\mathcal{F}(\mathcal{S}) = \{A \subseteq \Omega | \omega^* \in A\}.$$

Then  $\mathcal{F}(S)$  is the n-1-dimensional simplex dominated by  $\omega^*$ . Thus  $N = 2^{n-1}$ , and  $\mathbf{H}^{\mathbb{C}} := \mathbf{H}$  is the  $n \times 2^{n-1}$  matrix shown in Fig. 4, choosing the permutation of the  $\omega_i$  so that

$$\pi_i^* \ge \pi_{i+1}^*. \tag{5}$$

Thus from (3) we have

$$\vec{\pi}^{*T} = \mathbf{H}^{\mathbf{C}} \vec{m}^{T}, \tag{6}$$

where now  $\tilde{m}$  is a  $2^{n-1}$  vector.

Given  $\pi^*$  characterized in this way, we now wish to construct  $S^*$ . From (3), (4), and (6) we have

$$m_i^* = \pi_i^* - \pi_{i+1}^* = \mathbf{H}^{\mathbf{C}}(i)\vec{m}^T - \mathbf{H}^{\mathbf{C}}(i+1)\vec{m}^T$$

where  $\mathbf{H}^{\mathbf{C}}(n+1)$  is the zero vector  $(0, 0, \dots, 0)$ . We can thus construct a new  $n \times 2^{n-1}$  matrix  $\mathbf{H}^{\prime \mathbf{C}}$  such that

$$\bar{m}_n^{*T} = \mathbf{H}_{n \times 2^{n-1}}^{\prime \mathsf{C}} \, \bar{m}_{2^{n-1}}^T, \tag{7}$$

where  $\vec{m}^*$  is now an *n*, not a  $2^{n-1}$ , vector, and the subscripts indicate matrix dimensions. We do this by simply letting

$$\mathbf{H}^{\prime \mathbf{C}}(i) := \mathbf{H}^{\mathbf{C}}(i) - \mathbf{H}^{\mathbf{C}}(i+1).$$

The results are shown in Fig. 5.

## 5. Towards Measures of Distortion

We now wish to consider the distortion introduced to a consistent random set S by its consonant approximation  $S^*$ , and therefore the difference between Pl(A) and  $\Pi^*(A)$ .

#### 5.1. Towards a Measure

First, from Thm. (3) we know that  $\forall A \subseteq \Omega, \Pi^*(A) \leq \text{Pl}(A)$ . Therefore an appropriate measure for distortion could be something like

$$D(\mathcal{S}) := \sum_{A \subseteq \Omega} \operatorname{Pl}(A) - \Pi^*(A)$$

Since  $Pl(A) - \Pi^*(A) \in [0, 1]$ , therefore  $0 \leq D(S) \leq 2^{n-1}$ , and so it might be wise to consider  $\log_2 \qquad _2(D(S)) \leq n+1$ . We have yet to consider the circumstances under which D(S) = 1 so that  $\log_2(D(S)) = 0$  is a critical point.

But whichever measure might eventually be used, we need to characterize  $Pl(A) - \Pi^*(A)$  better. We do so in our linear algebraic formulation as follows. First, since S is consistent and  $S^*$  consonant, we need to define structures corresponding to the matrix  $\mathbf{G}^{\cap}$  restricted to those columns present. Thus define  $\mathbf{G}^{\cap S}$  as the matrix constructed by taking those columnns of  $\mathbf{G}^{\cap}$  corresponding to the  $2^{n-1}$ focal elements of S, and correspondingly  $\mathbf{G}^{\cap *}$  from those corresponding to the *n* focal elements of  $S^*$ . Then, using the subscripts on some elements to indicate their matrix dimensions, we have from (7)

$$\vec{\mathbf{P}}_{2^{n}-1}^{I} = \mathbf{G}_{2^{n}-1\times 2^{n-1}}^{OS} \vec{m}_{2^{n-1}}^{T} \vec{\Pi}_{2^{n}-1}^{*T} = \mathbf{G}_{2^{n}-1\times n}^{O*} \vec{m}_{n}^{*T} = \mathbf{G}_{2^{n}-1\times n}^{O*} \mathbf{H}_{n\times 2^{n-1}}^{C} \vec{m}_{2^{n-1}}^{T} (\vec{\mathbf{P}}_{1}^{I} - \vec{\Pi}^{*})_{2^{n}-1}^{T} = (\mathbf{G}^{OS} - \mathbf{G}^{O*} \mathbf{H}^{\prime C}) \vec{m}^{T}.$$

#### 5.2 An Example

The general case for n = 3 combined with a numerical example is shown in Fig. 6. The top of the figure shows m and Pl for a consistent random set on  $\Omega = \{x, y, z\}$ , where the  $\omega_i$  have been ordered according to (5). Then we have

$$\begin{aligned} \mathcal{F}(\mathcal{S}) &= \{\{x\}, \{x, y\}, \{x, z\}, \{x, y, z\}\} \\ \mathcal{F}(\mathcal{S}^*) &= \{\{x\}, \{x, y\}, \{x, y, z\}\} \\ \vec{m} &= \langle .1, .3, .2, .4 \rangle \end{aligned}$$

0-7803-7078-3/01/\$10.00 (C)2001 IEEE.

		$m_1$	$m_2$	$m_3$	$m_4$	$m_5$	$m_6$	$m_7$	$m_8$		$m_{2^{n-1}} - 2$	$m_{2^{n-1}} - 1$	$m_{2^{n-1}}$
<b>H</b> <sup>C</sup> =	$\omega_1$	/ 1	1	1	1	1	1	1	1		1	1	1
	$\omega_2$	0	1	0	1	0	1	0	1		1	0	1
	$\omega_3$	0	0	1	1	0	0	1	1		0	1	1
	$\omega_4$	0	0	0	0	1	1	1	1		1	1	1
	:	:	÷	÷	÷	÷	÷	÷	÷	۰.	÷	÷	:
	$\omega_n$	0	0	0	0	0	0	0	0		1	1	1 /

## Figure 4. $\mathbf{H}^{c}$ for a consistent randomset.

		$m_1$	$m_2$	$m_3$	$m_4$	$m_5$	$m_6$	$m_7$	$m_8$		$m_{2^{n-1}} - 2$	$m_{2^{n-1}} - 1$	$m_{2^{n-1}}$
H' <sup>C</sup> =	$\omega_1$	/ 1	0	1	0	1	0	1	0		0	1	0 \
	$\omega_2$	0	1	-1	0	0	1	-1	0		1	-1	0
	$\omega_3$	0	0	1	1	$^{-1}$	-1	0	0		-1	0	0
	$\omega_4$	0	0	0	0	1	1	1	1		0	0	0
	÷		÷	÷	÷	÷	÷	÷	÷	۰.	•	•	÷
	$\omega_n$	0 /	0	0	0	0	0	0	0		1	1	1 /

Figure 5.  $\mathbf{H}^{\prime c}$  for a consistent randomset.



Figure 6. (Top) A consistent random set, m and Pl. (Bottom left) The approximation  $m^*$ . (Bottom right) The approximation  $\Pi^*$ .

The final result was verified computationally, and verifies the results determined numerically from the alternate formulation

$$\Pi^*(A) = \bigvee_{\omega_i \in A} \pi_i^* = \bigvee_{\omega_i \in A} \mathbf{H}^{\mathsf{c}}(i) \vec{m}^T,$$

and shown in the figure.

## 6 Conclusion

The work presented here is, if only for space reasons, incomplete. In particular, the properties of D need to be explored more completely. Moreover, the linear algebraic formulation used here needs to be placed in a more general combinatorical context, involving Möbius [9, 10] and/or Fourrier transformations.

#### 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). I would like to thank Bill Oberkampf of Sandia in particular not only for his support of this work, but also for the technical conversations which prompted consideration of the question addressed here. I would also like to thank Leonid Gurvitz (Los Alamos National Laboratory) for his insights into better linear algebraic formulations of random set problems.

## References

- Dubois, Didier and Prade, Henri: (1990) "Consonant Approximations of Belief Functions", Int. J. Approximate Reasoning, v. 4, pp. 419-449
- [2] Joslyn, Cliff: (1994) Possibilistic Processes for Complex Systems Modeling, UMI Disseration Svcs, Ann Arbor MI, PhD dissertation
- [3] Joslyn, Cliff: (1994) "Possibilistic Approach to Qualitative Model-Based Diagnosis", Telematics and Informatics, v. 11:4, pp. 365-384
- [4] Joslyn, Cliff: (1996) "Aggregation and Completion of Random Sets with Distributional Fuzzy Measures", Int. J. of Uncertainty, Fuzziness, and Knowledge-Based Systems, v. 4:4, pp. 307-329
- [5] Joslyn, Cliff: (1997) "Measurement of Possibilistic Histograms from Interval Data", Int. J. General Systems, v. 26:1-2, pp. 9-33
- [6] 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
- [7] 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
- [8] Joslyn, Cliff: (1999) "Possibilistic Systems Theory Within a General Information Theory", in: 1999 Workshop on Imprecise Probabilities and their Applications, ed. G. de Cooman et al., pp. 206-215
- Kennes, Robert: (1992) "Computational Aspects of Möbius Transformations of Graphs", *IEEE Trans. on Systems, Man, and Cybernetics*, v. 22:2, pp. 201-223
- [10] Thoma, H Mathis: (1991) "Belief Function Computations", in: Conditional Logic in Expert Systems, ed. IR Goodman et al., pp. 269-308, North Holland