ADVANCES IN IMAGING AND ELECTRON PHYSICS, VOL. 148

# Interval and Fuzzy Analysis: A Unified Approach

WELDON A. LODWICK

*Department of Mathematical Sciences, University of Colorado at Denver and Health Sciences Center, Denver, Colorado 80217, USA*

I. Introduction . . . . . 76
A. Historical Background . . . . . 76
B. The Focus and Basic Themes . . . . . 79
1. Extension Principles . . . . . 80
2. Arithmetic . . . . . 81
3. Enclosure and Verification . . . . . 83
II. Interval Analysis . . . . . 85
A. Interval Extension Principle . . . . . 90
B. Interval Arithmetic . . . . . 93
1. Axiomatic Approach . . . . . 94
2. Interval Arithmetic from the United Extension: Constraint Interval Arithmetic . 95
3. Specialized Interval Arithmetic . . . . . 98
4. Comparison Between the Axiomatic and Extension Principle Approach to Interval Arithmetic . . . . . 109
C. Enclosure and Verification . . . . . 110
1. Enclosure of the Range of a Function . . . . . 113
2. Epsilon Inflation . . . . . 113
3. Defect Correction . . . . . 115
D. Algorithms and Software . . . . . 117
III. Fuzzy Set Theory . . . . . 119
A. Possibility and Necessity Distribution . . . . . 123
B. Semantics of Fuzzy Sets and Possibility and Necessity Distributions . . . . . 124
C. Fuzzy Extension Principles . . . . . 125
1. Fuzzy Arithmetic . . . . . 127
D. Enclosure and Verification . . . . . 129
IV. Analysis with Distributions . . . . . 129
A. Distribution Arithmetic . . . . . 131
1. Interval Convolution Methods . . . . . 132
2. Interval Histogram Methods . . . . . 136
3. Inverse Probability Method . . . . . 152
B. General Theory of Uncertainty . . . . . 165
1. Clouds . . . . . 165
2. Interval-Valued Probability Measures . . . . . 167
C. Generalized Extension Principles for Distribution . . . . . 184
References . . . . . 184



ISSN 1076-5670
DOI: 10.1016/S1076-5670(07)48002-8

# I. Introduction

## A. Historical Background

A unified approach to real-valued interval and fuzzy analysis emphasizing common themes is presented. Interval analysis and fuzzy analysis may be viewed as a bridge between deterministic problem solving and problems with generalized uncertainty (Zadeh, 2005). This presentation focuses on two key features common to both interval and fuzzy analysis: (1) the *extension principles* that generalize real-valued analysis-to-analysis on sets and functions (intervals, fuzzy sets, and possibility distributions), and (2) *verification* or *enclosure methods* that define lower and upper values and distributions with the view of obtaining approximate solutions to problems whose existence is verified and given within guaranteed bounds in an efficient and mathematically consistent way. The extension principles relate interval analysis and fuzzy analysis. The extension principles may point to how computation with numbers and sets is extended to computation with words. Enclosure methods are derived from the extension principles and indicate how interval, fuzzy, and possibilistic uncertainty can be incorporated, computed, and propagated in mathematical analysis on these entities. Enclosure methods, for example, in the substitution of interval arithmetic operations for every algebraic operation in a continuous algebraic function, obtain lower and upper bounds on the result. Knowing how to incorporate, compute, and propagate uncertainty in mathematical analysis problem solving is key to the use of interval, fuzzy, and possibilistic entities in practice. Enclosure methods are one of the most important contributions of interval analysis and fuzzy analysis to mathematical analysis. One emphasis of this chapter is enclosure. While not explicitly derived, the enclosure methods developed here may equally be applied to probabilistic distributions.

Professor Lothar Collatz, in a lecture titled *Monotonicity and Optimization*, given to the Mathematics Department at Oregon State University, August 5, 1974, stated, "The idea of *ordering* is more fundamental than *distance* because from ordering we can get distance but not the other way around" (author notes). The real number system contains within itself the most fundamental mathematical order to which both interval analysis and fuzzy analysis relate. Real-valued arithmetic and derived distance generate mathematical analysis. If we take Professor Collatz's insight seriously, any new structure—interval analysis or fuzzy analysis in particular—has as its first task the elucidation of order and subsequently derived distance (measure) from that order. Since both interval numbers and fuzzy numbers, as will be seen from their definition, relate themselves to *sets* of real numbers and graphs in $\mathbb{R}^2$, respectively, the order needs to be derived from that associated with sets and distributions.

Moreover, since interval numbers and fuzzy sets model uncertainty on one hand and amplification and flexibility on the other, it is clear that the idea of order and the derived distance (measure) generated will be flexible; that is, they require choices. The choice that needs to be made depends on the semantics of the problem to a greater extent than in the deterministic setting. Our approach takes as its point of departure arithmetic, which has as its underlying the order of *sets* of real numbers. Subsethood is what will be taken as the order on sets. Thus, since sets are the primary entity, the way to handle set-valued functions whose domains are interval numbers, fuzzy sets, and possibility distributions is a major task of this chapter.

Functions and algorithms are the mathematical embodiment of cause and effect. A mathematical function can be considered the symbolic expression of the essence of science, the study of what ties cause to effect. Thus, functions are central to scientific endeavors. Intervals and fuzzy sets are relatively new entities of mathematical study; therefore, the most useful approaches to their analyses are still being uncovered. Nevertheless, functions are such a central part of mathematics that they must be confronted by any developing field; for this reason, they are the theme of this presentation. As will be demonstrated, functions as defined by the extension principle are a key feature to the interconnection between interval analysis and fuzzy analysis. Extension principles and verification methods express how to obtain functions. Moreover, knowing how to relate interval analysis to fuzzy analysis clears the way to application and amplification of results in each discipline, as well as the creation of new ones. This monograph contends that the extension principles are the bridge between the two disciplines, and enclosure methods apply the extension principles to obtain lower and upper bounds on computations with interval, fuzzy, and possibilistic entities to possibly guarantee that the lower and upper bounds that are obtained contain within their range the correct result. Researchers have searched for efficient methods to compute with intervals, fuzzy sets, and possibility distributions. Methods to compute lower and upper values (numerical lower and upper bounds on intervals or lower and upper distributions on fuzzy sets and possibility distributions) that are developed and presented use a min/max calculus (presented below) to attain efficiency in computation.

Some material found here was first developed and presented by Dubois and Prade (1980, 1987a, 1987b, 1991), and later with colleagues Kerre and Mesiar (Dubois *et al.*, 2000b). For example, the interrelationship between fuzzy set theory and interval analysis was first discussed in 1984 by Dubois and Prade (1980, page 58) and later in their 1984 technical report (Dubois and Prade, 1987b). The central theme of this presentation is the extension principle and enclosure and verification as unifying concepts in interval and fuzzy analysis. Dubois and Prade give prominence to the extension principle. However, they

do not deal explicitly with enclosure and verification methods, especially as arising from the extension principle. Since the extension principle is the analysis of function over sets, this discussion begins with set-valued functions and shows how these directly tie into Moore's united extension and Zadeh's extension principle. This forms the basis for constrained interval arithmetic (Lodwick, 1999), as well as for the ensuing analyses.

This chapter is intended for applied professionals, researchers, and students who have some familiarity with fuzzy set theory or interval analysis and wish to study the synergisms between the two fields. Some basic knowledge of interval analysis (as found in Moore, 1979) and fuzzy set theory (as found in Klir and Yuan, 1995) is assumed. For example, it is assumed that the reader has understanding of an $\alpha$*–cut* of a fuzzy membership function.

Louis Rall, at the beginning of his talk at *SIAM Workshop on Validated Computing* (May 22–25, 2002, Toronto, Canada), said (I am paraphrasing), "My early career can be characterized by what I learned from R.E. Moore. My later career can be characterized by what I further learned from R.E. Moore." It is certainly true that this chapter reflects what I learned from R.E. Moore, D. Dubois, and H. Prade as I was beginning to learn interval analysis in graduate school and fuzzy set theory in my early academic career. Subsequently, this chapter reflects what I further learned from R.E. Moore, D. Dubois, and H. Prade.

Interval analysis and fuzzy set theory, as active fields of research and application, are relatively new mathematical disciplines receiving the impetus that defined them as separate fields of study in 1959 and 1965, respectively, with R.E. Moore's technical reports on interval analysis and his doctoral dissertation (Moore, 1959a, 1962; Moore and Yang, 1959b; Moore *et al.*, 1960) and Zadeh's seminal papers on fuzzy set theory (Zadeh, 1965, 1968, 1975, 1978). The connection between interval analysis and possibility theory [possibility theory historically arose out of fuzzy set theory (Zadeh, 1978)] is evident in the mathematics of uncertainty. The theory of interval analysis models, among other things, the uncertainty arising from numerical computation, which can be considered a source of ambiguity. Fuzzy set theory and possibility theory model, among other things, the uncertainty of vagueness and ambiguity arising from the transitional nature of entities and a lack of information, respectively. This chapter will clarify the distinction between fuzzy set theory and possibility theory. Since intervals can be considered a particular type of fuzzy set (as we will see), one view is that fuzzy set theory is a more general theory than interval analysis. That is, interval analysis can be thought of as dealing with a particular *type* of uncertainty whose general theory is described by fuzzy sets. However, interval analysis developed as part of the then emergent field of numerical analysis initially had three directions: (1) *computational error analysis* (automatically computed error including

rounding); (2) *verified computing*, which Moore first called *range arithmetic*, subsequently called interval arithmetic [later Aberth (1988) developed a separate arithmetic he called range arithmetic, which differs from Moore's use of these words]; and (3) the derivation of the underlying algebraic structure of floating-point numbers called *computer algebra*. Fuzzy sets developed in at least two directions of interest to this presentation: (1) *possibility theory*, and (2) *fuzzy interval analysis*. These have direct linkages with what has become known as *generalized theory of uncertainty* (see Zadeh, 2005). Although the two fields can be considered as having a common root, interval analysis and fuzzy set theory are independent fields whose cross-fertilization has been a relatively recent phenomenon (Dubois and Prade, 1991; Dubois *et al.*, 2000b; Lodwick and Jamison, 2003a; Lodwick, 2002).

Of note, *all* the interval analysis and fuzzy analysis in this monograph is over sets of real numbers. All sets are over real numbers, that is, real-valued intervals and real-valued distributions (in the case of fuzzy membership functions, possibilistic distributions, and probabilistic distributions). Moreover, when the word *box* is used in the context of intervals, it is understood that if the analysis is in $\mathbb{R}$, the box refers to an interval $[a, b]$. In the context of $\mathbb{R}^n$, a box is a rectangular $n$-dimensional hyperrectangle $[a_1, b_1] \times \cdots \times [a_n, b_n]$, $a_i, b_i \in \mathbb{R}, i = 1, \ldots, n$.

### *B. The Focus and Basic Themes*

The thesis of this presentation in that an underlying mathematical theory from which interval and fuzzy analysis can be viewed and understood is the theory of set functions particularized to intervals and fuzzy sets and associated lower and upper approximations of the resultants. Therefore, this development focuses on two areas common to both interval and fuzzy analysis: the *extension principles* and *enclosure*.

Functions are a central part of mathematics and science, so their application to interval and fuzzy sets is a crucial step in the development of both fields. The extension principles of R.E. Moore (Moore *et al.*, 1960) and L. Zadeh (Zadeh, 1965) are directly related to an earlier development and more general extension principle (Strother, 1952, 1955). A relatively recent treatment of set-valued functions is Audin and Frankkowska (1990). Of course, set-valued functions extend real-valued functions to functions on intervals and fuzzy sets. The extension principle used in interval analysis is called the *united extension* (Moore *et al.*, 1960; Moore, 1962). In fuzzy set theory, it is called simply the *extension principle* (Zadeh, 1965). Since arithmetic operations are continuous real-valued functions, excluding division by zero, the extension principles can be used to define interval and fuzzy arithmetic.

*Enclosure* means approximations that produce lower and upper values (an interval or a functional envelope, depending on the context) to the theoretical solution, which lies between the lower and upper values or functions. Efficient methods to compute lower and upper approximations are desired and necessary in practice. When enclosure is part of mathematical problem solving, it is called *verification*, formally defined in the following text.

The point of view that the extension principle is an important thread, which can be used to relate and understand the various principles of uncertainty that are of interest to this discussion (interval, fuzzy, and possibility), leads to a direct relationship between the associated arithmetic. Moreover, lower and upper approximation pairs for fuzzy sets allow for simpler computation using min/max calculus and lead to enclosures with careful implementation. Arithmetic is a gateway to mathematical analysis. The extension principle is how arithmetic is defined for uncertainty entities of interest, and thus the extension principle provides a path to understand mathematical analysis on these entities.

Three basic themes ensue: (1) *extension principles*, (2) *arithmetic* derived from axioms and extension principles, and (3) *enclosure* and *verification*. The sections after this introduction consider the three basic themes in the context of interval analysis (Section II), fuzzy set theory (Section III), and distributions (Section IV).

## 1. *Extension Principles*

The extension principle is key because it defines how real-valued expressions are represented in the context of intervals, fuzzy sets, distributions, and generalized uncertainty. The extension principle can be viewed as one of the main unifying concepts between interval analysis and fuzzy set theory. Moreover, the extension principle is one way to define arithmetic on intervals, fuzzy sets, and distributions.

All extension principles associated with intervals and fuzzy sets can be considered as originating from set-value mappings or graphs. Generally, an extension principle defines how to obtain functions whose domains are sets. Accomplishing this for real numbers is clear. It is more complex for sets because how to obtain resultant well-defined entities must be defined.

Set-valued maps have a very long history in mathematics. Relatively recently, Strother's 1952 doctoral dissertation (Strother, 1952) and two subsequent papers (Strother, 1955, 1958) define the *united extension* for set-valued functions for domains possessing specific topological structures. R.E. Moore applied Strother's united extension to intervals. In doing so, as will be seen in Section II, Moore had to show that the topological structures on intervals were among those that Strother developed. Having done this, he retained the

name *united extension* as the extension principle particularized to intervals. In fact, Strother co-authored the technical report that first used the set-valued extension principle on intervals. That is, Moore's united extension (the interval extension principle) is a set-valued function whose domain is the set of intervals, and the range is an interval for those underlying functions that are continuous.

Zadeh's extension principle (Zadeh, 1965) explains how functions of fuzzy sets are derived from real-valued functions. It describes, among other things, how to compute with fuzzy sets. That is, Zadeh's extension principle can be thought of as a set-valued function where the domain elements are fuzzy sets, and the range values are fuzzy sets for the appropriate maps, called *membership functions* (defined in Section III). The extension principle was generalized and made more specific to what are now called *fuzzy numbers* or *fuzzy intervals* by various researchers beginning with Nguyen (1978). The details are found in Section III.

### 2. *Arithmetic*

Interval arithmetic is central to fuzzy arithmetic and can be derived axiomatically or from R.E. Moore's *united extension*. Of special interest is the latter approach, especially in deriving a constrained interval arithmetic in Section II, which will have implications for fuzzy arithmetic. Moore (Moore, 1959a, 1962; Moore and Yang, 1959b; Moore *et al.*, 1960) developed interval analysis arising from Moore and Lodwick (2003)

> ... the observation that if a real number $A$ is computed, and a rigorous bound $B$ on the total error in $A$ as an approximation to some unknown number $X$, that is, $|X - A| \leq B$, then no matter how $A$ and $B$ are computed, it is known for certain that $X$ lies in the interval $[A-B, A+B]$.

Computations with intervals, especially in conjunction with computer implementations, arose naturally. In fact, Moore's initial work is tied directly to numerical error analysis.

There are two direct precursors of Moore's development of interval analysis in 1959: Warmus (1956) and Sunaga (1958). Moore's initial work references and extends in significant ways Sunaga's work. Moore develops computational methods, incorporates computer rounding, develops for the first time automatic numerical error analysis (gets the computer to calculate roundoff, numerical truncation, and numerical method error estimations), and extends interval arithmetic to interval analysis.

An interval can be considered as a *set* and a *number*. On the real number line, with the usual meaning of the order relation $\leq$, an interval $[a, b]$ is the set of all real numbers $\{x \mid a \leq x \leq b\}$. As a number, an interval $X$ is a

pair of numbers $\{a, b\}$, the left and right end points of the interval. Analysis on intervals, since intervals are sets, requires set-valued functions, limits, integration, and differentiation theory. This is done via the *united extension* (Moore, 1966).

The axioms of interval arithmetic as articulated by Warmus (1956), Sunaga (1958), and Moore (1959a) are as follows. It is noted that while Warmus' notation is different, the operations are the same.

***Definition 1.*** For all arithmetic functions $\circ \in \{+, -, \div, \times\}$, $[x] \circ [y] = \{x \circ y \mid x \in [x] \text{ and } y \in [y]\}$ where $[x]$ is an arbitrary interval $X$.

In particular, we have the following:

1. **Addition**:

$$[a, b] + [c, d] = [a + c, b + d]. \tag{1}$$

2. **Subtraction**:

$$[a, b] - [c, d] = [a - d, b - c]. \tag{2}$$

3. **Multiplication**:

$$[a, b] \times [c, d] = \big[\min\{ac, ad, bc, bd\}, \max\{ac, ad, bc, bd\}\big]. \tag{3}$$

4. **Division**:

$$[a, b] \div [c, d] = [a, b] \times [1/d, 1/c] \quad \text{where } 0 \notin [c, d]. \tag{4}$$

There is an extended interval arithmetic that incorporates the case where $0 \in [c, d]$ for division (Hansen, 1975; Kahan, 1968b). Moreover, there are various ways to approach interval arithmetic; for example, see Dempster (1974), Neumaier (1993), Nickel (1969), and Stolfi *et al.* (1994). In fuzzy arithmetic, the axioms of interval arithmetic apply to each $\alpha$-*cut* of a fuzzy set membership function as long as the entity is a *fuzzy number* or *fuzzy interval* (defined in Section III).

***Definition 2.*** An interval arithmetic based on axioms 1–4 above is called *axiomatic interval arithmetic*.

***Remark.*** The axioms 1–4 above essentially compute the maximum and minimum values of the set $\{z = x \circ y \mid x \in [a, b], y \in [c, d]\}$, where $\circ \in \{+, -, \times, \div\}$ and the two intervals $[a, b]$, $[c, d]$ are considered independent. Thus, axiomatic interval arithmetic is a type of min/max calculus since the values arising from axiomatic interval arithmetic only involve the end points (min/max) of the intervals.

The implementation of interval arithmetic on the computer, for which the goal is to account for all errors, including numerical and truncation errors, is called *rounded interval arithmetic* (Moore, 1959a; Moore and Yang, 1959b). Kulisch (1971) and Kulisch and Miranker (1981, 1986) axiomatized rounded interval arithmetic and uncovered the resultant algebraic structure, called a *ringoid*. While specialized extended languages (Pascal-XSC, C-XSC) and chips were developed for interval and rounded interval data types, incorporating the ideas set forth by Moore, Kulisch, and Miranker (among other researchers), the most successful rounded interval tool is undoubtedly *INTLAB*, a downloadable software package that runs in conjunction with *MATLAB* with imbedded interval arithmetic, rounded interval arithmetic, and some interval analytic methods, in particular computational linear algebraic methods.

### 3. *Enclosure and Verification*

Archimedes' (Archimedes of Siracusa, 1897; Phillips, 1981) approach to the computation of the circumference of a circle using outer circumscribed and inner inscribed regular polygons whose perimeter is a straightforward calculation is an enclosure and verification method, perhaps the first one. The essential part of enclosure and verification is that a solution is mathematically proven to exist (perhaps is unique) and lies between the *computed* lower and upper bound values (real numbers for our purposes). That is, verification guarantees that the solution exists (and perhaps is unique) in a mathematical sense. Enclosure consists of lower and upper bounds containing the solution. Verification in the case of Archimedes' computation of the circumference of a circle is the geometrical fact (theorem) that the perimeter of the circumscribed regular polygon is greater than the circumference of a circle and that the inscribed regular polygon has a perimeter less than that of the circumference of a circle. Often fixed-point theorems (contractive mapping theorems, for example) are used to verify the existence of solutions in mathematical analysis. These theorems are used from the point of view of verification of hypotheses so that interval analysis may be used to mathematically calculate guaranteed bounds (for example, compute a Lipschitz condition, accounting for *all* numerical and truncation errors, that is less than 1, for example), which if held means that the mapping is contractive (hence, a *computed* solution exists). The methods to compute lower and upper bounds in a mathematically correct way on a computer must account for numerical and computer truncation error. This is one of the core research areas of interval mathematics. One of the primary applications of interval analysis consists of enclosures for verification. As will be seen, the equivalent for fuzzy sets and possibility theory is the computation of functional envelopes

(interval-valued probability or clouds). Interval verification methods obtain interval enclosures containing the solution(s) within their bounds. In interval analysis, verification means that existence is mathematically demonstrated, and valid bounds on solutions are given. When possible and/or relevant, uniqueness is mathematically determined. Thus, verification in the context of a computational process that uses intervals for a given problem means that a solution, say $x$, is verified (mathematically) to *exist*, and the computed solution is returned with lower and upper *bounds*, $a$ and $b$, such that the solution, shown to exist, is guaranteed to lie between those bounds, that is, $a \leq x \leq b$. *Uniqueness* is determined when possible or desirable.

Although not often thought of in these terms, possibility and necessity pairs, when constructed from an underlying though perhaps unknown distribution according to Jamison and Lodwick (2002), envelope the distribution. As such, they are functional enclosures. Thus, enclosure in the context of distributions is understood to be the construction of lower and upper functions, $g(x)$ and $h(x)$, to a given function $f(x)$ such that $g(x) \leq f(x) \leq h(x)$. This is desired not only when $x$ is a real number or vector, but also when $x$ consists of vectors of distributions such as of random variables, intervals, fuzzy sets, and/or possibilities. This is an especially difficult problem when $f(x)$ is such a complex expression. The problem of computing such lower and upper bounds for various types of domains whose elements are described by a variety of uncertainty distributions is important in problems dealing with risk analysis, optimization under uncertainty, and simulations when inputs are distributions.

Fuzzy set membership functions only give information about upper values of uncertainty, as noted by Dubois, Moral, and Prade (Dubois *et al.*, 1997). The lower distribution for a fuzzy set is the *x-axis*. That is, in the context of enclosures, a fuzzy entity is enclosed between $g(x) = 0$ and $f(x) = h(x)$, where $h(x)$ is the membership function. The same is not true with a probability density function $f(x)$, since a probability density function is considered "deterministic" in the sense that it captures the uncertainty with "certainty," that is, $g(x) = f(x) \leq f(x) \leq h(x) = f(x)$. For fuzzy sets, it is possible to obtain a tighter bound on the lower end, $g(x)$, which can be accomplished by constructing probability-based possibility and necessity pairs via interval analysis (Berleant and Goodman-Strauss, 1998; Olsen, 2005; Tonon, 2004), with probability-based possibility and necessity pairs (Jamison and Lodwick, 2002), or with clouds (Neumaier, 2004a). These are discussed later.

This exposition begins with interval analysis, its early development, and its extension principle, which led to interval arithmetic and interval analysis. The various interval arithmetics that arise from axioms, mathematical analysis, and the direct application of the extension principle all attempt to solve the outstanding problems of dependency, tight bounds, and implementability. Enclosure and verification methods, which are presented next, are applied

to problems in mathematical analysis. Enclosure and verification methods demonstrate the power and usefulness of interval analysis. Fuzzy set theory as it relates to the basic themes of the extension principle, arithmetic, analysis, enclosure, and verification methods follow next. The concluding section discusses enclosure and verification methods for uncertainty distributions under the subsection "Distribution Arithmetic."

## II. Interval Analysis

Interval analysis has an early history in Archimedes' computation of the circumference of a circle (Archimedes of Siracusa, 1897). Interval analysis as developed by Moore (Moore, 1959a; Moore and Yang, 1959b; Moore *et al.*, 1960) arose from the attempt to compute the error bounds of the numerical solutions of a finite state machine (computer), for which the roundoff error was automatically accounted for by the computer itself. This led to the investigation of computations with intervals as the entity, data type, that enabled automatic error analysis. Hansen writes (Hansen, 2001):

> R.E. Moore (Moore, 1999) states that he conceived of interval arithmetic and some of its ramifications in the spring of 1958. By January of 1959, he had published (Moore, 1959a) a report on how interval arithmetic could be implemented on a computer. A 1959 report (Moore and Yang, 1959b) showed that interval computations could bound the range of rational functions and integrals of rational functions. Theoretical and practical interval arithmetic were differentiated. Reference (Moore, 1965) discusses interval-valued functions, interval contractions, a metric topology for interval numbers, interval integrals, and contains an extensive discussion of Moore's use of interval analysis to bound the solution of ordinary differential equations.

Although there are five known direct and clear precursors to Moore's version of interval arithmetic and interval analysis beginning in 1924 (Burkill, 1924; Dwayer, 1951; Fischer, 1958; Sunaga, 1958; Warmus, 1956; Young, 1931), it was Moore who worked out rounded computer arithmetic and fully developed interval analysis, that is, the continuous mathematics of intervals. Moore (1999) describes his *independent* discovery of interval analysis. He states (personal communication, December 22, 2006):

> When I submitted the first draft of my thesis to George Forsythe at Stanford, in late 1959 … Forsythe insisted that I do a careful literature search for prior work. I spent the next several months in the mathematics library at Stanford, and thereby found the works of Burkhill, Dwyer, Fischer, Sunaga and R.C. Young.

Moore was the first, in 1959, to use interval methods in the solution methods of ordinary differential equations. While it can be argued that Burkill (1924) was the first to deal with interval analysis since his article considers functions of intervals (which includes arithmetic since the algebraic operations are functions), Burkill's interest is *functionals* of intervals since he maps intervals to real numbers such as occur in derivatives and integration. Intervals as entities in and of themselves, for example, functions that map intervals to intervals, were not the focus of Burkill's work. In 1931 Young developed the arithmetic for sets of numbers and was interested in properties of limits in this more general setting. Moreover, Young worked out interval arithmetic and the commutative, associative, and distributive law of scalars (real numbers) over intervals. In 1951 Dwyer particularized Young's work to compact sets of numbers (intervals). Both Warmus (1956) and Sunaga (1958) had the full development of axiomatic interval arithmetic. Sunaga recognized the importance of interval arithmetic in computational mathematics but did not proceed further. In 1958 Fischer reported a computer program that uses two computer words that automate propagated and roundoff error. One computer word holds the approximate value of the variable; the other word holds the value representing the bound of previous computations and roundoff errors of the approximation. The computer program includes the bound contained in the second word to compute subsequent bounds resulting from the processing being done. Notwithstanding this earlier development, Moore and his colleagues are responsible for developing the early theory, extensions, vision, a wide array of applications, and the actual implementation of interval analysis to computers. Moore's major contributions include at least the following:

1. Moore recognized how to use intervals in computational mathematics, now called *numerical analysis*.
2. Moore extended and implemented the arithmetic of intervals to computers.
3. Moore's work was influential in creating Institute of Electrical and Electronics Engineers (IEEE) standards for accessing computer's rounding processes, which is a necessary step in obtaining computer-generated verified computations (Kulisch and Miranker, 1986).
4. Moore developed the *analysis* associated with intervals where the united extension plays a key role in achieving this. One such major achievement was to show that Taylor series methods for solving differential equations are not only more tractable, but more accurate (Moore, 1979).
5. Moore was the first to recognize the usefulness of interval analysis in computer verification methods, especially for solutions to nonlinear equations using the interval Newton method in which the method includes verification of existence and uniqueness of solution(s) by successive intersection of intervals, which is the key to successful interval Newton

> method approaches. Sunaga (1958) did have a version of the Newton method. However, it was Moore who recognized that intersections of resulting intervals were key to an interval Newton method. "In fact George Forsythe [R.E. Moore's major professor] asked an important question. ... He asked, 'what about an interval version of Newton's method?' I set to work on finding one, and almost at once saw that an intersection of interval Newton steps was *necessary* in order to obtain other very important properties which the interval Newton method then had." (R.E. Moore, personal communication, December 22, 2006).

The term *range arithmetic*, which was the name Moore used for what is now called interval arithmetic, was only used in Moore (1959a). While Moore's earliest articles point to the use of interval arithmetic to compute the range of a function (hence, the original name of range arithmetic, and so to global optimization), it was Hansen who was the first to develop interval global arithmetic. Walster states how interval global optimization was first conceived:

> Compute upper bounds on (a function) $f$ at various sampled points. The smallest computed upper bound, $\bar{f}$, is an upper bound for $f^*$ (the global minimum). Then delete subboxes of the current box wherein $f > \bar{f}$. The concept of generating and using an upper bound in this way was introduced by Hansen (1980). [Personal communication among Moore, William Walster, Hansen, and Lodwick; January 10, 2007].

An interval $[a, b]$ is defined as the set of all real numbers $X = \{x\colon\ a \leq x \leq b\}$ on the real number line with the usual meaning of the order relation $\leq$. A second definition of an interval is as a pair of real numbers that are the end points of the interval. Thus, from the point of view of interval analysis, intervals on the real line have a dual nature, as sets of real numbers and as a new kind of number represented by pairs of real numbers, with an arithmetic, interval arithmetic, developed axiomatically, (1)–(4), and consistent with the set interpretation. The logic associated with interval analysis is one of certain containment. The sum of two intervals certainly contains the sums of all pairs of real numbers, one from each of the intervals. This follows from the definitions of interval arithmetic based on simple properties of the order relation $\leq$ on the real line.

Intersections and unions of intervals also have an algebra and are computed in a straightforward manner. These definitions, unlike those of fuzzy set theory found in the sequel, come from classical set theory. Intersections and unions are crucial in defining what is meant by solutions to simultaneous equations and inequalities, as well as the fundamental building blocks of logical statements. For a given interval $[a, b]$ and a given real number $x$, the statement $x \in [a, b]$ is either true or false. There is no vagueness or ambiguity,

except for roundoff when the statement is implemented on a computer. For two intervals $A_1$ and $A_2$, if we know that $x \in A_1$ and $x \in A_2$, then we also know with certainty that $x \in A_1 \cap A_2$. These statements have certainty (except possibly for implementations on the computer where roundoff error can become a factor), unlike statements of this type in fuzzy set theory. Interval arithmetic and the interval analysis developed from it do not assign any measure of possibility or probability to parts of an interval. A number $x$ is either in an interval $A$ or it is not. By introducing probability distributions or possibility distributions on an interval, and using level sets, integrals, or other measures, a connection between intervals and fuzzy sets can be made.

The application of functions to intervals is accomplished through what R.E. Moore calls the *united extension* (an extension principle) already used in Moore and Yang (1959b) and formally defined in Moore *et al.* (1960), particularized to intervals from the more general extension theory for set-valued functions developed in Strother (1952). In particular, using interval computations, the range of values of a real-valued continuous function $f$ of a single real variable $x$ is contained in an interval $[c, d]$ when $x \in [a, b]$ where $f([a, b]) \subseteq [c, d]$. Such an interval exists since $f$ is continuous and $[a, b]$ is a compact set so that $f$ attains both a minimum and maximum (and all values in between). Thus, if $f$ is continuous, we may be interested in finding an interval containing a 0 of $f$ that is a solution to the equation $f(x) = 0$. If $0 \notin [c, d]$, for which we can test using $0 < c$ or $d < 0$, then we know that there is no solution in $[a, b]$. On the other hand, if $0 \in [c, d]$, then it is *possible* that $f$ has a 0 in $[a, b]$. It is not certain because it is in the nature of interval computation that it cannot generally find exact ranges of values. That is, we may have $0 \in [c, d]$ but $0 \notin f([a, b])$. Techniques have been developed over the past four decades for reducing overestimation of interval ranges of mappings to any prescribed tolerance with enough computing, thus enabling the analysis of whether there exists a 0 of the function in the interval.

Computer implementations of intervals add another dimension of uncertainty to intervals. If the lower limit of the range of a function is an irrational number (or even a particular rational number), then it cannot be represented exactly with floating-point numbers, which are a subset of the rational numbers. However, we can find a close rational approximation where "close" is associated with the computer's precision. Thus, it is necessary when computing with intervals to round down (left) lower end points and round up (right) upper end points, and so compute a result that contains the set of all possible results.

An interval number $U = [\underline{u}, \bar{u}]$ considered as a fuzzy set (defined in Section III), has a membership function

$$\mu_U(x) = \begin{cases} 1 & \text{for } \underline{u} \le x \le \bar{u}, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, interval analysis may be considered a subset of fuzzy set theory. As a probability distribution, various authors have considered an interval number as one of two probability density functions.

1. The distribution
$$p(x) = \begin{cases} \frac{1}{\bar{u}-\underline{u}} & \text{for } \underline{u} \le x \le \bar{u}, \underline{u} < \bar{u}, \\ 0 & \text{otherwise.} \end{cases}$$
2. An interval may represent the fact that all we know is the support so that the set of all distributions, $p(x)$, with support $\text{supp}_{p(x)} = [\underline{u}, \bar{u}]$ is in the interval. Many, including Berleant (1993), take this point of view.

***Remark.*** The issue of how to interpret intervals as probabilities is emphasized in Dubois *et al.* (1997), where semantics is also discussed. Moreover, it turns out that intervals form the underlying fundamental principle of fuzzy and interval numbers (Dubois and Prade, 2005; Fortin *et al.*, 2006) (discussed in Section III).

Neumaier (1990, page 1) states, "Interval arithmetic is an elegant tool for practical work with inequalities, approximate numbers, error bounds, and more generally with certain convex and bounded sets." He goes on to say that intervals arise naturally in:

1. Physical measurements
2. Truncation error, the representation of an infinite process by a finite process
   (a) Representation of numbers by finite expansions
   (b) Finite representation of limits and iterations
3. Numerical approximations
4. Verification of monotonicity and convexity
5. Verification of the hypotheses of fixed-point theorems (the contraction mapping theorem or Brouwer's fixed-point theorem are examples)
6. Sensitivity analysis, especially as applied to robotics
7. Tolerance problems

Interval arithmetic and analytic methods have been used to solve an impressive array of problems given that these methods capture error (modeling, roundoff, and truncation) so that rigorous accounting of error together with the contraction mapping theorem or the Brouwer fixed-point theorem allow for computer verification of existence, uniqueness, and enclosure. In particular, Tucker (2002), using interval analysis, solved a long-outstanding problem [Smale's 14th conjecture (Smale, 1998)] by showing that the Lorenz equations do possess a strange attractor. Davies (2005, page 1352) observes,

> Controlled numerical calculations are also playing an essential role as intrinsic parts of papers in various areas of pure mathematics. In some areas

> of nonlinear PDE, rigorous computer-assisted proofs of the existence of solutions have been provided.... These use *interval arithmetic* to control the rounding errors in calculations that are conceptually completely conventional.

Another long-standing problem, Kepler's conjecture about the densest arrangement of spheres in space, was solved by Hales (2000) using interval arithmetic. Ten problems were posed by Trefethen in the January/February, 2002 issue *SIAM News*; each had a real number solution, and the objective was to obtain a 10-digit solution to each of the problems. The book of Bornemann *et al.* (2004) documents not only the correct solutions, but the analysis behind the problems. One of the authors, Wagon (personal communication), indicated that,

> Intervals were extremely useful in several spots. Problem 4: Designing an optimization algorithm to solve it by subdividing.... Problem 2: Intervals could be used to solve it by using smaller and smaller starting interval until success is reached. Moreover, Problems 2, 4, 7, and 9, intervals yield proofs that the digits are correct.

Chapter 4 of Bornemann *et al.* (2004) contains an exposition of interval optimization. Robust stability analysis for robots that uses interval analysis methods, performed with the aid of a computer that is verified to be mathematically correct, can be found in Daney *et al.* (2004), Jaulin (2001a), or Jaulin *et al.* (2001b). There are excellent introductions to interval analysis beginning with Moore's book (Moore, 1966) (also see other texts listed in the references). A more recent introduction can be found in Corliss (2004) and downloaded from http://www.eng.mu.edu/corlissg/PARA04/READ_ME.html. Moreover, introductions can be downloaded from the interval analysis website (http://www.cs.utep.edu/interval-comp).

### *A. Interval Extension Principle*

Moore recognizes, in three Lockheed Aircraft Corporation technical reports (Moore, 1959a; Moore and Yang, 1959b; Moore *et al.*, 1960), that the extension principle is a key concept. Interval arithmetic, rounded interval arithmetic, and computing range of functions can be derived from interval extensions. Of issue is how to compute ranges of set-valued functions. This requires continuity and compactness over interval functions, which in turn needs well-defined extension principles.

In 1960, Moore (Moore *et al.*, 1960) used for the first time in an explicit way, the extension principle for intervals called the *united extension*, which particularizes set-valued extensions to sets that are intervals. Following

Strother's development (Strother, 1952, 1955), a specific topological space is the starting point.

***Definition 3.*** A topological space $\{X, \Omega\}$ on which open sets are defined ($X$ is a set of points, and $\Omega$ is family of subsets of $X$) has the property that $X \in \Omega$, $\emptyset \in \Omega$, finite intersections, and uncountable unions of sets in $\Omega$ are back in $\Omega$. The topological space $\{X, \Omega\}$ is called a $\mathbf{T}_1$-*space* if for every $x \in X$ and $y \in X$ (distinct points of $X$), there is an open set $O_y$ containing $y$ but not $x$. Metric spaces are $\mathbf{T}_1$. In addition, if there exists an open set $O_x$ containing $x$ such that $O_x \cap O_y = \emptyset$, then the space is called a $\mathbf{T}_2$-*space*, or a Hausdorff space. The set of subsets of $X$ is denoted by $S(X)$.

**Lemma 1** *(Strother, 1955). If $f : X \to Y$ is a continuous multivalued function and $Y$ is a* $\mathbf{T_1}$-space*, then $f$ is closed for all $x$. That is, the image of a closed set in $X$ is closed in $Y$.*

***Remark.*** This means that $f$ is a point-closed function. In other words, when the sets of $X$, $S(X)$, are retracted to points, then the set-valued function $F$ on $S(X)$ becomes the original function $f$ when defined through the retract. Phrased in another manner, if $f$ is a real-valued function, then the mapping extended to sets of $X \subseteq \mathbb{R}$, $S(X)$, where the range is endowed with a $\mathbf{T}_1$ topology, is well defined.

***Remark.*** (Moore *et al.*, 1960). If $f : X \to Y$ is an arbitrary mapping from an arbitrary set $X$ into an arbitrary set $Y$, the united extension of $f$ to $S(X)$, denoted $F$, is defined as follows.

$$F : S(X) \to S(Y) \quad \text{where}$$
$$F(A) = \big\{ f(a) \mid \forall a \in A, A \in S(X) \big\}, \quad \text{in particular}$$
$$F\big(\{x\}\big) = \big\{ f(x) \mid x \in \{x\}, \{x\} \in S(X) \big\} \quad \text{where } \{x\} \text{ is a singleton set.}$$

Thus,

$$F(A) = \bigcup_{a \in A} f(a).$$

This definition, as we shall see, is quite similar to the fuzzy extension principle of Zadeh, where the union is replaced by the supremum.

**Theorem II.1** *(Strother, 1955). Let $X$ and $Y$ be compact Hausdorff spaces and $f : X \to Y$ continuous. Then the united extension of $f$, $F$, is continuous. Moreover, $F$ is closed.*

There are fixed "point" (set) theorems associated with these set-valued maps, the united extensions, that can be found in Strother (1955) and not repeated here since we are not interested in the full generality of these theorems, but only as they are related to intervals. Moore (Moore *et al.*, 1960, 1962, 1966) particularized the ideas of Strother (1952, 1955) to spaces consisting of subsets of $\mathbb{R}$ that are closed and bounded (real intervals), which we denote as $S([\mathbb{R}])$, where $S([X])$ denotes the set of all intervals on any set of real numbers $X \subseteq \mathbb{R}$. To this end, Moore needed to develop a topology on $S([\mathbb{R}])$ that proved to be Hausdorff. The results of interest associated with the united extension for intervals are the following (Moore, 1966):

1. *Isotone Property*: A mapping $f$ from a partially ordered set $(X, r_X)$ into another $(Y, r_Y)$, where $r_X$ and $r_Y$ are relations, is called *isotone* if $x r_X y$ implies $f(x) r_Y f(y)$. In particular, the united extension is isotone with respect to intervals and the relation $\subseteq$. That is, for $A, B \in S([X])$, if $A \subseteq B$, then $F(A) \subseteq F(B)$.
2. *The Knaster–Tarski Theorem*: An isotone mapping of a complete lattice into itself has at least one fixed point.

The Knaster–Tarski theorem implies that $F$, the united extension, $F : S([\mathbb{R}]) \to S([\mathbb{R}])$, has at least one fixed "point" (set) in $S([\mathbb{R}])$, which may be the empty set, and has an important numerical consequence. Consider the sequence $\{X_n\}$ in $S(X)$ defined by (choosing $X_0 = X$)

$$X_{n+1} = F(X_n).$$

Since

$$X_1 \subseteq F(X_0) = F(X) \subseteq X = X_0,$$

then, by induction,

$$X_{n+1} \subseteq X_n.$$

Let

$$Y = \bigcap_{n=0}^{\infty} X_n.$$

The following is true (Moore *et al.*, 1960). If $x = f(x)$ is any fixed point of $f$ in $X$, then $x \in X_n$ for all $n = 0, 1, 2, \ldots$ so that $x \in Y$ and

$$x \in F(Y) \subseteq Y.$$

Thus, $X_n$, $Y$, and $F(Y)$ contain all the fixed points $f$ in $X$. If $Y$ and/or $F(Y)$ is empty, then there are no fixed points of $f$ in $X$. Newton's method is a fixed-point method, so that the above theorem pertains to a large class of problems. Moreover, these enclosures lead to computationally verified solutions when implemented on a computer with rounded interval arithmetic.

### *B. Interval Arithmetic*

Interval arithmetic was defined axiomatically by Young (1931), Dwayer (1951), Warmus (1956) in 1956, and then independently by Sunaga (1958). Moore (1959a, 1959b) rediscovers and extends interval arithmetic to rounded interval arithmetic, thereby allowing interval arithmetic to be useful in computational mathematics. There are two approaches to interval arithmetic. The first is that obtained by application of the united extension. The second approach is the axiomatic approach. An interval arithmetic and associated semantics allow for "intervals," $[a, b]$, for which $a > b$ (Gardeñes *et al.*, 1986; Hanss, 2005). This arithmetic is related to directed interval arithmetic (see Section II.B.3) and has some interesting applications to fuzzy control (Bondia *et al.*, 2006; Sáinz, 2001).

The basic axioms associated with interval arithmetic are (1)–(4). They are more fully developed in Moore (1966). There are various properties associated with the axiomatic approach to interval arithmetic that differ from those of real numbers and the constraint interval arithmetic defined subsequently. In particular, interval arithmetic derived from axioms is subdistributive. Thus, from Moore (1979) we have for intervals $X$, $Y$, and $Z$:

1. $X + (Y + Z) = (X + Y) + Z$ The associative law for addition
2. $X \cdot (Y \cdot Z) = (X \cdot Y) \cdot Z$ The associative law for multiplication
3. $X + Y = Y + X$ The commutative law for addition
4. $X \cdot Y = Y \cdot X$ The commutative law for multiplication
5. $[0, 0] + X = X + [0, 0] = X$ Additive identity
6. $[1, 1] \cdot X = X \cdot [1, 1] = X$ Multiplicative identity
7. $X \cdot (Y + Z) \subseteq X \cdot Y + X \cdot Z$ The subdistributive property.

***Example.*** Moore (1979, page 13) points out that

$$[1, 2](1 - 1) = [1, 2](0) = 0,$$

whereas

$$[1, 2](1) + [1, 2](-1) = [-1, 1].$$

***Remark.*** From Moore's (Moore, 1959a) implementation of Sunaga (1958) [neither Moore nor Sunaga seem to have been aware of Warmus' earlier work (Warmus, 1956)], he states that

$$X \circ Y = \{z \mid z = x \circ y, x \in X, y \in Y, \circ \in \{+, -, \times, \div\}\},$$

which means that Moore applies the united extension for distinct intervals $X$ and $Y$. However, Moore abandons this united extension definition and

develops axioms. The axioms lead to a simplification of the operations since one need not account for multiple occurrences, while at the same time it leads to overestimation (which is severe at times). From the beginning, Moore was aware of the problems of overestimation associated with multiple occurrences of the same variable in an expression. Moreover, it is apparent that, from the axiomatic approach, $X - X$ is *never* 0 unless $X$ is a real number (a zero width interval). Moreover, $X \div X$ is *never* 1 unless $X$ is a real number (a zero width interval).

## 1. *Axiomatic Approach*

The axiomatic approach to interval arithmetic considers all instantiations of variables as independent. That is, the Young, Warmus, Sunaga, and Moore axiomatic approach to interval arithmetic is one in which multiple occurrences of a variable in an expression are considered as independent variables. While axiomatic interval arithmetic is quite simple to implement, it leads to overestimations.

***Example.*** Consider

$$f(x) = x(x-1), \quad x \in [0,1].$$

Using the axiomatic approach,

$$[0,1]\big([0,1]-1\big) = [0,1][-1,0] = [-1,0]. \tag{5}$$

However, the interval containing $f(x) = x(x-1)$ is $[-0.25, 0]$. This is because the two instantiations of the variable $x$ are taken as independent when they are dependent. The united extension $F(A)$, which is

$$F\big([0,1]\big) = \bigcup_{x\in[0,1]} \big\{f(x)\big\} = [-0.25, 0],$$

was not used.

If the calculation were $x(y-1)$ for $x \in [0,1]$, $y \in [0,1]$, then the tightest interval containing $x(y-1)$, its united extension, indeed is $[-1,0]$. Note that the subdistributivity property does not use the united extension in computing $X \cdot Y + X \cdot Z$, but instead considers $X \cdot Y + W \cdot Z$, where $W = X$. Partitioning the interval variables (that are repeated) leads to closer approximation to the united extension. That is, take the example above and partition the interval in which $x$ lies.

***Example.*** Consider $x(x-1)$ again, but $x \in [0, 0.5] \cup [0.5, 1]$. This yields

$$[0, 0.5]\big([0, 0.5] - 1\big) \cup [0.5, 1]\big([0.5, 1] - 1\big) \tag{6}$$

$$= [0, 0.5][-1, -0.5] \cup [0.5, 1][-0.5, 0] \tag{7}$$

$$= [-0.5, 0] \cup [-0.5, 0] = [-0.5, 0], \tag{8}$$

which has an overestimation of 0.25 compared with an overestimation of 0.5 when the full interval [0, 1] was used.

In fact, for operations that are continuous functions, a reduction in width leads to estimations that are closer to the united extension and in the limit, to the exact united extension value (Moore, 1962, 1966, 1979; Neumaier, 1990). Other approaches that find ways to reduce the overestimation arising from the axiomatic approach have proved extremely useful; these include the centered, mean value, and slope forms (Hansen, 1992; Kearfott, 1996a; Moore, 1979; Neumaier, 1990, 2004b; Ratschek and Rokne, 1988). More recently, the Taylor models discussed below (Makino and Berz, 2003, 2005) exhibit a high order of convergence and minimization of the dependency illustrated in the above two examples.

## 2. *Interval Arithmetic from the United Extension: Constraint Interval Arithmetic*

The power of the axiomatic approach to interval arithmetic is its application simplicity. Its complexity is at most four times that of real-valued arithmetic. However, the axiomatic approach to interval arithmetic leads to overestimations in general because it takes every instantiation of the same variable independently. As seen below, the united extension, when applied to sets of real numbers, is global optimization, which generally is *NP-Hard*. Conversely, simple notions such as

$$X - X = 0 \tag{9}$$

and

$$X \div X = 1, 0 \notin X \tag{10}$$

are desirable properties and can be maintained if the united extension is used to define interval arithmetic [as will be seen below and in Lodwick (1999)]. In the context of fuzzy arithmetic, which uses interval arithmetic, Klir (1997) looked at fuzzy arithmetic, which was constrained to account for Eqs. (9) and (10) from a case-based approach. What is given next was developed in Lodwick (1999) independently of Klir (1997) and is more general than the case-based method. Constraint interval arithmetic is

derived directly from the united extension rather than axiomatically or case-based.

It is known that applying interval arithmetic to the union of intervals of decreasing width yields tighter bounds on the result that converges to the united extension interval result (Moore, 1962). Of course, for $n$-dimensional problems, "intervals" are rectangular parallelepipeds (boxes), and as the diameters of these boxes approach 0, the union of the result approaches the correct bound for the expression. Partitioning each of the sides of the $n$-dimensional box in half has complexity of $O(2^n)$ for each split. Theorems proving convergence to the exact bound of the expression and the rates associated with the subdivision of intervals can be found in Hansen (1992), Kearfott (1996a), Moore (1979), Neumaier (1990, 2004b) or Ratschek and Rokne (1988). What is proposed here is to redefine interval numbers in such a way that dependencies are explicitly kept. The ensuing arithmetic will be called *constraint interval arithmetic*. This new arithmetic is the derivation of arithmetic directly from the united extension of Strother (1952).

An interval number is redefined (Lodwick, 1999) into an equivalent form next as the *graph* of a function of one variable and two coefficients or parameters.

***Definition 4.*** An interval $[\underline{x}, \bar{x}]$ is the *graph* of the real single-valued function $X^I(\lambda_x)$, where

$$X^I(\lambda_x) = \lambda_x \underline{x} + (1 - \lambda_x)\bar{x}, \quad 0 \leq \lambda_x \leq 1. \tag{11}$$

Strictly speaking, in Eq. (11), since the numbers $\underline{x}$ and $\bar{x}$ are known (inputs), they are *coefficients*, whereas $\lambda_x$ is varying, although constrained between 0 and 1, hence the name *constraint interval arithmetic*. Note that Eq. (11) defines a *set* representation explicitly, and the ensuing arithmetic is developed on sets of numbers. The algebraic operations are defined as follows:

$$\begin{aligned} Z = [\underline{z}, \bar{z}] &= X \circ Y \\ &= \{z \mid z = x \circ y, \forall x \in X^I(\lambda_x), y \in Y^I(\lambda_y), 0 \leq \lambda_x, \lambda_y \leq 1\} \\ &= \{z \mid z = (\lambda_x \underline{x} + (1 - \lambda_x)\bar{x}) \\ &\qquad \circ (\lambda_y \underline{y} + (1 - \lambda_y)\bar{y}), 0 \leq \lambda_x \leq 1, 0 \leq \lambda_y \leq 1\}, \end{aligned} \tag{12}$$

where

$$\underline{z} = \min\{z\}, \quad \bar{z} = \max\{z\}, \quad \text{and} \quad \circ \in \{+, -, \times, \div\}. \tag{13}$$

***Remark.*** It is clear from Eq. (13) that constraint interval arithmetic requires a global optimization.

When the operations use the same interval, no exceptions are necessary, as in Klir (1997). Using only Eq. (12), we obtain,

$$\begin{aligned} Z = [\underline{z}, \bar{z}] &= X \circ X \\ &= \{z \mid z = (\lambda_x \underline{x} + (1-\lambda_x)\bar{x}) \circ (\lambda_x \underline{x} + (1-\lambda_x)\bar{x}), 0 \le \lambda_x \le 1\} \\ &= \{z \mid z = (\lambda_x \underline{x} + (1-\lambda_x)\bar{x}) \circ (\lambda_x \underline{x} + (1-\lambda_x)\bar{x}), 0 \le \lambda_x \le 1\}. \end{aligned} \tag{14}$$

This results in the following properties:

1. Addition of the same interval variable:

$$\begin{aligned} X + X &= \{z \mid z = (\lambda_x \underline{x} + (1-\lambda_x)\bar{x}) \\ &\qquad + (\lambda_x \underline{x} + (1-\lambda_x)\bar{x}), 0 \le \lambda_x \le 1\} \\ &= \{z \mid z = 2(\lambda_x \underline{x} + (1-\lambda_x)\bar{x}), 0 \le \lambda_x \le 1\} = [2\underline{x}, 2\bar{x}]. \end{aligned}$$

2. Subtraction of the same interval variable:

$$\begin{aligned} X - X &= \{z \mid z = (\lambda_x \underline{x} + (1-\lambda_x)\bar{x}) \\ &\qquad - (\lambda_x \underline{x} + (1-\lambda_x)\bar{x}), 0 \le \lambda_x \le 1\} = 0. \end{aligned}$$

3. Division of the same interval variable, $0 \notin X$:

$$\begin{aligned} X \div X &= \{z \mid z = (\lambda_x \underline{x} + (1-\lambda_x)\bar{x}) \\ &\qquad \div (\lambda_x \underline{x} + (1-\lambda_x)\bar{x}), 0 \le \lambda_x \le 1\} = 1. \end{aligned}$$

4. Multiplication of the same interval variable with $\underline{x} < \bar{x}$

$$\begin{aligned} X \times X &= \{z \mid z = (\lambda_x \underline{x} + (1-\lambda_x)\bar{x}) \\ &\qquad \times (\lambda_x \underline{x} + (1-\lambda_x)\bar{x}), 0 \le \lambda_x \le 1\} \\ &= \{z \mid z = (\lambda_x^2 \underline{x}^2 + 2(1-\lambda_x)\underline{x}\lambda_x\bar{x} + (1-\lambda_x)^2\bar{x}^2), 0 \le \lambda_x \le 1\} \\ &= [\min\{\underline{x}^2, \bar{x}^2, 0\}, \max\{\underline{x}^2, \bar{x}^2, 0\}]. \end{aligned}$$

To verify that this is the interval solution, note that as a function of the single variable $\lambda_x$, the product, $X \times X$, is

$$f(\lambda_x) = (\bar{x} - \underline{x})^2 \lambda_x^2 + 2\underline{x}(\bar{x} - \underline{x})\lambda_x + \underline{x}^2,$$

which has a critical point at

$$\lambda_x = -\frac{\underline{x}}{\bar{x} - \underline{x}}.$$

Thus,

$$\underline{z} = \min\left\{ f(0),\, f(1),\, f\left(-\frac{\underline{x}}{\bar{x} - \underline{x}}\right)\right\},$$

$$\bar{z} = \max\left\{ f(0),\, f(1),\, f\left(-\frac{\underline{x}}{\bar{x} - \underline{x}}\right)\right\},$$

$$\underline{z} = \min\{\underline{x}^2, \bar{x}^2, 0\}, \qquad \bar{z} = \max\{\underline{x}^2, \bar{x}^2, 0\}$$

as is obvious. Of course, if $\underline{x} = \bar{x}$, then $X \times X = x^2$.

5. $X(Y + Z) = XY + XZ$.

Constraint interval arithmetic is the complete implementation of the united extension, and it provides an algebra that possesses an additive inverse, a multiplicative inverse, and a distributive law.

## 3. *Specialized Interval Arithmetic*

Various interval arithmetic approaches have been developed in addition to the axiomatic and united extension approaches. Different representations of intervals have been created and include the development of range and rational arithmetic. These purport to simplify operations and/or obtain more accurate results using arithmetic. Another issue addressed by researchers was how to extend interval arithmetic, called *extended interval arithmetic*, to handle unbounded intervals that may be entered or result from a division by 0. The general space of improper intervals, which includes extended interval arithmetic, called *directed interval arithmetic*, was developed subsequently. Next, *generalized interval arithmetic* and its more recent generalizations, *affine arithmetic* and *Taylor model arithmetic*, deals with the problem of reducing overestimation that characterizes the axiomatic approach to interval arithmetic. *Triplex arithmetic* and its generalization, *quantile arithmetic*, were developed to carry more information than an interval carries. These specialized arithmetics are presented next.

**a. Interval Arithmetic with Different Representation.** Two representations of numbers can be used in interval arithmetic. Range arithmetic is the midpoint/error form of representing an interval. Rational arithmetic as it applies to interval arithmetic is mentioned because of its potential "speed" per unit of work and accuracy as a way to represent floating point numbers. Therefore, its representation would propagate less overestimation.

*Range Arithmetic.* Range arithmetic was developed by Aberth (1988, pages 13–25). A range number [Eq. (15)] is really an interval where

$$X = m \pm \epsilon = [\underline{x}, \bar{x}], \quad m = \frac{\underline{x} + \bar{x}}{2}, \epsilon = \frac{\bar{x} - \underline{x}}{2}, \tag{15}$$

and the arithmetic (Aberth, 1988), page 15, is given by:

$$X_1 + X_2 = (m_1 \pm \epsilon_1) + (m_2 \pm \epsilon_2) = (m_1 + m_2) \pm (\epsilon_1 + \epsilon_2), \quad (16)$$

$$X_1 - X_2 = (m_1 \pm \epsilon_1) - (m_2 \pm \epsilon_2) = (m_1 - m_2) \pm (\epsilon_1 + \epsilon_2), \quad (17)$$

$$\begin{aligned} X_1 \times X_2 &= (m_1 \pm \epsilon_1) \times (m_2 \pm \epsilon_2) \\ &\subseteq (m_1 m_2) \pm \big(\epsilon_1 |m_2| + \epsilon_2 |m_1| + \epsilon_1 \epsilon_2\big), \end{aligned} \quad (18)$$

$$\begin{aligned} X_1 \div X_2 &= (m_1 \pm \epsilon_1) \div (m_2 \pm \epsilon_2) \\ &\subseteq \left(\frac{m_1}{m_2}\right) \pm \left(\frac{\epsilon_1 + |\frac{m_1}{m_2}|\epsilon_2}{|m_2| - \epsilon_2}\right). \end{aligned} \quad (19)$$

Using range arithmetic, $X - X = [-2\epsilon, 2\epsilon]$, and $X \div X = 1 \pm \frac{2\epsilon}{|m|-\epsilon}$, which means that the problem of repeated variables still remains, although with improved bounds in general.

*Rational Arithmetic.* Rational arithmetic is arithmetic on numbers whose representations are fractions. This is particularly useful since rational numbers have simple continued fraction representations. Moreover, there are rational arithmetic chips that use continued fraction representation of numbers (Korenerup and Matula, 1983). This is, in general, a more precise representation of floating-point numbers and computer arithmetic, although this approach has not been an accepted part of computer hardware. Aberth (1978, 1988, Chapter 5) shows how to use rational arithmetic to form more accurate representations for range (interval) arithmetic. To provide a computer interval arithmetic, one simply rounds outwardly all data and all intermediate rational arithmetic calculations. Since the round-down and round-up values are rational, this process is quite simple. In general, the result will be more accurate per available bytes devoted to floating-point number.

**b. Interval Arithmetic on the Extended Real Number System.** Interval arithmetic carried out on the computer must deal with division by 0 (among other challenges). Therefore, the axioms are extended to include arithmetic on values that are infinite so that they must deal with the extended real number system. Two approaches have been developed—extended real interval arithmetic and directed interval arithmetic.

*Extended Real Interval Arithmetic.* Interval arithmetic on the set of extended real numbers is called *extended interval arithmetic*. Kahan (1968b) was the first to propose the extension to handle, among other things, division by intervals containing 0 and subsequent processing of the resulting intervals. This is especially useful in Newton's method, and in this context Hansen (1978) uses interval arithmetic over the extended real numbers. A review of the extended real interval system can be found in Walster (1998). *INTLAB*,

C-XSC, PASCAL-XSC, Sun Microsystems Fortran, and Mathematica support extended interval arithmetic. The most recent approach to extended interval arithmetic is due to Pryce and Corliss (2006), who develop the theory and implementation of containment sets for interval arithmetic. The idea is to consider intervals as sets of real numbers where, in the implementation, interval sets become abstract data types. The arithmetic becomes operations on sets.

*Directed Interval Arithmetic.* Two approaches to extended interval arithmetic were developed. Kahan's approach dealt with the problem of how to incorporate plus and minus infinity as end points of intervals in interval arithmetic. The approach of Ortolf (1969) and Kaucher (1973, 1980) was to mathematically complete the set of intervals $I(\mathbb{R})$ to its closure $\overline{I(\mathbb{R})}$ by including what they call *nonregular intervals*. However, previously, Warmus (1956, 1961) had considered this space and its arithmetic. Alefeld and Herzberger (1983, page 8) state:

> These intervals are interpreted as intervals with negative width. The point intervals $[a, a]$ are no longer minimal elements with respect to the ordering $\subseteq$. All the structures of $I(\mathbb{R})$ are carried over to $I(\mathbb{R}) \cup \overline{I(\mathbb{R})}$, and a completion through two improper elements $p$ and $-p$ is achieved. In this manner the division by an interval $A = [\underline{a}, \bar{a}]$ with $\underline{a} \leq 0 \leq \bar{a}$, $\underline{a} \neq \bar{a}$, can also be defined.

This approach was studied by Gardeñes *et al.* (1986) and Hanss (2005). Popova (1998) states,

> Directed interval arithmetic is obtained as an extension of the set of normal intervals by improper intervals and a corresponding extension of the definitions of the interval arithmetic operations. The corresponding extended interval arithmetic structure possesses group properties with respect to addition and multiplication operations and a number of other advantages.

**c. Interval Arithmetic—Reducing the Effects of Dependencies.** Moore recognized from the beginning the problems associated with dependencies and the axiomatic approach. Various approaches have addressed this issue. Three of these are discussed: generalized interval arithmetic, affine arithmetic, and Taylor model arithmetic.

*Generalized Interval Arithmetic.*

1. One of the first attempts to deal with the fact that axiomatic interval arithmetic often yields results that are not sharp (overestimates) is generalized

interval arithmetic (Hansen, 1975). An interval is represented by

$$X = [\underline{x}, \bar{x}] = y + [-c, c], \quad c \geq 0,$$
$$y - c = \underline{x} \quad \text{and} \quad y + c = \bar{x}. \tag{20}$$

An arbitrary point $x \in X$ is expressed as

$$x = y + \xi, \quad \xi \in [-c, c]. \tag{21}$$

Observe that Eq. (21) is precisely the range arithmetic representation [Eq. (15)] published 9 years after generalized interval arithmetic. Whereas range arithmetic is an arithmetic on intervals, generalized interval arithmetic is an interval arithmetic that specifically deals with reducing the effects of dependencies. Range arithmetic does not deal explicitly with the issue of dependency. This means that any interval $X_k$ $(k > n)$, depending of the previous intervals $j = 1, \ldots, n$, has the form

$$X_k = Y_k + \sum_{l=1}^{n} \xi_l Z_{lk}, \tag{22}$$

where $Y_k$ and $Z_{lk}$ are previously computed intervals $k = 1, 2, \ldots, l = 1, 2, \ldots, n$, and $\xi_l \in [-c_l, c_l]$. An interval $X_k$ of the form Eq. (22) is called a *generalized interval*. The actual interval is

$$X_k = Y_k + \sum_{l=1}^{n} [-c_l, c_l] Z_{lk}$$
$$= Y_k + [-1, 1] \sum_{l=1}^{n} c_l z_{lk}, \tag{23}$$

where $z_{lk} = |Z_{lk}| = \max\{|\underline{Z}_{lk}|, |\overline{Z}_{lk}|\} = \max_{z \in Z_{lk}} |z|$. The general arithmetic operations are defined in Hansen (1975) as follows:

2. Generalized interval: **Addition/Subtraction**

$$X_k = X_i \pm X_j$$
$$= Y_i \pm Y_j + \sum_{l=1}^{n} \xi_l (Z_{li} \pm Z_{lj}).$$

Let $Y_k = Y_i \pm Y_j$ and $Z_{lk} = Z_{li} \pm Z_{lj}$, then

$$X_k = Y_k + \sum_{l=1}^{n} \xi_l Z_{lk},$$

which is again in the form of generalized interval arithmetic.

3. Generalized interval: **Multiplication**

$$X_k = Y_i Y_j + \sum_{l=1}^{n} \xi_l (Y_i Z_{li} + Y_j Z_{lj}) + \sum_{l=1}^{n} \sum_{m=1}^{n} \xi_l \xi_m Z_{li} Z_{lj}. \tag{24}$$

For $l = m$, replace $[-c_l, c_l]^2$ by $[0, c_l^2]$, and for $l \neq m$ replace $\xi_l \xi_m$ by $\xi_l[-c_m, c_m]$, which means that Eq. (24) can be replaced by

$$X_k = Y_k + \sum_{l=1}^{n} \xi_l Z_{lk}, \tag{25}$$

where

$$Y_k = Y_i Y_j + \sum_{l=1}^{n} [0, c_l^2] Z_{li} Z_{lj},$$

$$Z_{lk} = Y_i Z_{lj} + Y_j Z_{li} + \sum_{m=1, l\neq m}^{n} [-c_m, c_m] Z_{mj}$$

$$= Y_i Z_{lj} + Y_j Z_{li} + [-1, 1] z_{il} \sum_{m=1, l\neq m}^{n} c_m z_{jm}$$

and $z_{il} = |Z_{il}|$, $z_{jm} = |Z_{jm}|$. The result of multiplication of a generalized interval [Eq. (25)] is again a generalized interval.

4. Generalized interval: **Division**

$$X_k = \frac{X_i}{X_j} = \left( Y_i + \sum_{l=1}^{n} \xi_l Z_{li} \right) \Bigg/ \left( Y_j + \sum_{m=1}^{n} \xi_m Z_{mj} \right)$$

$$= \frac{Y_i}{Y_j} + \frac{\sum_{l=1}^{n} \xi_l (Y_j Z_{li} - Y_i Z_{lj})}{Y_j (Y_j + \sum_{m=1}^{n} \xi_m Z_{mj})}$$

$$= Y_k + \sum_{l=1}^{n} \xi_l Z_{lk},$$

where $Y_k = \frac{Y_i}{Y_j}$, $z_{mj} = |Z_{mj}|$,

$$Z_{lk} = \frac{\xi_l (Y_j Z_{li} - Y_i Z_{lj})}{Y_j (Y_j + [-1, 1] \sum_{m=1}^{n} c_m z_{mj})},$$

and $\xi_m$ is replaced by $[-c_m, c_m]$.

Generalized interval arithmetic has been successfully applied to problems associated with computer graphics of equations and surface rendering (Tupper, 1996).

*Affine Arithmetic.* Another more recent approach to minimize the effects of overestimation due to dependencies is affine arithmetic (Stolfi *et al.*, 1994). A number $x$, whose value is subject to uncertainty, has the representation

$$x = x_0 + x_1\epsilon_1 + \cdots + x_n\epsilon_n, \tag{26}$$

where $x_i$ are coefficients (known, real) and the $\epsilon_i \in [-1, 1]$. That is, $x_i$ represents the magnitude of "error" and $\epsilon_i$ represents the $i$th uncertainty that is contributing to the total uncertainty represented by the interval. To recover an interval from the affine representation of a number [Eq. (26)], each of the $\epsilon_i$ is replaced by $[-1, 1]$ so that

$$X = [x_0 - \xi, x_0 + \xi], \tag{27}$$

where $\xi = \sum_{i=1}^{n} |x_i|$. Clearly, if one is given an interval $X = [\underline{x}, \bar{x}]$, to obtain the affine number representation in Eq. (26),

$$x = x_0 + x_1\epsilon_1,$$

where $x_0 = (\underline{x} + \bar{x})/2$, $x_2 = (\underline{x} - \bar{x})/2$, and $\epsilon_1 \in [-1, 1]$, where the subscript is used to distinguish the variable $x_0$ from all other variables.

***Example.*** Suppose $x = 4 + 3\epsilon_1 + 2\epsilon_2 + \epsilon_3$, and $y = 2 + \epsilon_1 - 3\epsilon_2 + \epsilon_4$. From this representation, we see that $x$ depends on variables 1, 2, and 3, whereas $y$ depends on variables 1, 2, and 4. Note that the dependency information is carried forward, where $X = [-2, 10]$, and $Y = [-3, 7]$. Interval analysis, which does not carry forward dependencies, would obtain a sum of

$$Z = X + Y = [-2, 10] + [-3, 7] = [-5, 17].$$

Affine arithmetic yields

$$z = x + y = 6 + 4\epsilon_1 - \epsilon_2 + \epsilon_3 + \epsilon_4,$$

from which we obtain $Z = [-1, 13]$.

Addition and subtraction of affine numbers is straightforward. However, the challenge comes with multiplication and division in such a way that the interval result is guaranteed to enclose the true result. Moreover, obtaining an affine number for arbitrary transcendental functions, and composites of these, is another challenge. These have been computed for interval arithmetic in such

a way that roundoff errors are incorporated in the final result. The challenge in affine arithmetic as it is in quantile, range, and generalized interval arithmetic, is how to implement multiplication and division. These matters can be found in Stolfi *et al.* (1994).

**Multiplication** (Stolfi *et al.*, 1994)

$$\begin{aligned} z &= xy \\ &= \left(x_0 + \sum_{i=1}^{n} x_i\epsilon_i\right)\left(y_0 + \sum_{i=1}^{n} y_i\epsilon_i\right) \\ &= x_0y_0 + \sum_{i=1}^{n}(x_0y_i + y_0x_i)\epsilon_i + \left(\sum_{i=1}^{n} x_i\epsilon_i\right)\left(\sum_{i=1}^{n} y_i\epsilon_i\right) \qquad (28) \\ &= z_0 + \sum_{i=1}^{n} z_i\epsilon_i + Q(\epsilon_1, \ldots, \epsilon_n), \qquad (29) \end{aligned}$$

where $z_i = x_0y_i + y_0x_i$, and

$$\begin{aligned} Q(\epsilon_1, \ldots, \epsilon_n) &= \left(\sum_{i=1}^{n} x_i\epsilon_i\right)\left(\sum_{i=1}^{n} y_i\epsilon_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_j\epsilon_i\epsilon_j \qquad (30) \\ &\subseteq \sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_j[-1, 1][-1, 1] \qquad (31) \\ &= [-1, 1]\sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_j \qquad (32) \\ &= [\underline{q}, \bar{q}]. \qquad (33) \end{aligned}$$

To obtain an affine representation of Eq. (28) that contains the product,

$$\begin{aligned} z &= \hat{z}_0 + \sum_{i=1}^{n} z_i\epsilon_i + \frac{\underline{q} + \bar{q}}{2} + \frac{\underline{q} - \bar{q}}{2}\epsilon_{n+1} \\ &= z_0 + \sum_{i=1}^{n+1} z_i\epsilon_i, \end{aligned}$$

where $z_0 = \hat{z}_0 + (\underline{q} + \bar{q})/2$, and $z_{n+1} = (\underline{q} - \bar{q})/2$. Stolfi *et al.* (1994) go on to successfully and efficiently apply affine arithmetic to computer graphics.

*Taylor Model Arithmetic.* The Taylor model (Makino and Berz, 2003, 2005) is perhaps the most successful tractable modern approach to deal with dependencies. The Taylor model is a method to do arithmetic on functions.

Since arithmetic of numbers is a function, these methods can be applied to ordinary interval arithmetic. For the general Taylor model, Makino and Berz (2003) state:

> ...the Taylor model has the following fundamental properties:
> 1. The ability to provide enclosures of any function with a finite computer code list by a Taylor polynomial and a remainder bound with a sharpness that scales with order $(n + 1)$ of the width of the domain.
> 2. The ability to alleviate the dependency problem in the calculation.
> 3. The ability to scale favorable to higher-dimensional problems.

The basic definition is:

***Definition 5.*** Let $f : D \subseteq \mathbb{R}^k \to \mathbb{R}$ be a function that is $(n + 1)$ times continuously differentiable on an open set containing the domain $D$. Let $x_0 \in D$ and $P$ the $n$th-order Taylor polynomial of $f$ around $x_0$. Let $I$ be an interval such that

$$f(x) \in P(x - x_0) + I \quad \text{for all } x \in D. \tag{34}$$

Then the pair $(P, I)$ is called an $n$th-*order Taylor model of* $f$ *around* $x_0$ *on* $D$.

Let $T_1 = (P_1, I_1)$ and $T_2 = (P_2, I_2)$. Then, the *Taylor model arithmetic* is given as follows.

1. **Addition** (see Makino and Berz, 2003, page 384): $T_1 + T_2 = (P_1, I_1) + (P_2, I_2) = (P_1 + P_2, I_1 + I_2)$.
2. **Multiplication** (see Makino and Berz, 2003, page 384): $T_1 \times T_2 = (P_1, I_1) \times (P_2, I_2) = (P_{1\cdot 2}, I_{1\cdot 2})$, where $P_{1\cdot 2}$ is the part of the polynomial $P_1 \cdot P_2$ up to order $n$. The interval part is

$$I_{1\cdot 2} = B(P_e) + B(P_1) \cdot I_2 + B(P_2) \cdot I_1 + I_1 \cdot I_2,$$

where $P_e$ is the part of the polynomial $P_1 \cdot P_2$ of orders $n + 1$ up to $2n$, and $B(P)$ denotes a bound of $P$ on the domain $D$. Makino and Berz go on to state that $B(P)$ is required to be at least as "sharp" as direct interval evaluation of $P(x - x_0)$ on $D$. The rules for subtraction and division are clear.

***Remark.*** When the polynomials are simply numbers, $P_1 = x$, $P_2 = y$, $I_1$, and $I_2$ represent the interval bounds on the roundoff error. As such, the Taylor model for numbers resembles Fischer's (1958) approach.

***Remark.*** When evaluating functions and doing arithmetic or analysis with functions, it is clear that the Taylor model is an excellent approach since

overestimation on dependencies is lessened, if not eliminated, up to the precision of the floating-point representation being provided the computer program code list is of reasonable size. It is also clear that the Taylor model is one that enclosures results and thus is able to verify (formally defined in the sequel).

A downloadable software package for the Taylor model is available at http://www.beamtheory.nscl.msu.edu/cosy.

**d. Interval Arithmetic—The Carrying of More Uncertainty Information.** Intervals carry only the bound information on the uncertainty they represent. Two methods of carrying more uncertainty information, short of doing arithmetic on distributions (Section IV), are presented next. Triplex arithmetic carries along a "central" value, whereas quantile arithmetic (a generalization of triplex arithmetic) carries an arbitrary but finite amount of intermediate information about the distribution of the uncertainty that lies between the end points of the interval.

*Triplex Arithmetic.* Triplex arithmetic (Nickel, 1969) is a way to carry more information about the uncertainty beyond the bounds that are represented by the end points of the interval (the end points of the support if it is a distribution) by keeping track of a *main value* within the interval in addition to its end points. According to Nickel (1969), triplex arithmetic started as a project initiated in 1966 at the University of Karlsruhe to develop a compiler and to demonstrate its usefulness for solutions to problems in numerical analysis. Three-valued set theory has also been studied by Klaua (1969) and Jahn (1980). The presentation here is a synopsis of Nickel (1969).

A triplex number is $X = [\underline{x}, \tilde{x}, \bar{x}]$, $\underline{x} \le \tilde{x} \le \bar{x}$, where $[\underline{x}, \bar{x}]$ is the interval, and $\tilde{x}$ is called the *main value*. The main value could be an average value if this is known. The arithmetic is straightforward in the sense that

$$Z = X \circ Y = [\underline{z}, \tilde{z}, \bar{z}] \quad \text{where } \circ \in \{+, -, \times, \div\}, \tag{35}$$

$$[\underline{z}, \bar{z}] = [\underline{x}, \bar{x}] \circ [\underline{y}, \bar{y}] \quad \text{(obtained from interval arithmetic)}, \tag{36}$$

$$\tilde{z} = \tilde{x} \circ \tilde{y}. \tag{37}$$

In statistical arithmetic, if the main value is to be interpreted as something like a mean or mode, an issue arises when the $\circ$ in Eq. (37) is multiplication or division. Thus, the interpretation of the resulting main value from a statistical point of view using Eq. (37), is problematic. That is, the semantic value of the resulting "main value" in multiplication and division, as computed by Eq. (37), is not clear.

*Quantile Arithmetic.* Quantile arithmetic (Dempster, 1969, 1974) is a way to carry more information about the uncertainty bound in a probabilistic and statistically faithful way than triplex arithmetic. While it is more complex (as will be seen), it does have a well-defined probabilistic and statistical semantics. In fact, triplex arithmetic can be represented by quantile arithmetic. In particular, quantile arithmetic approximates distributions whose support is an interval (which can be infinite for extended interval arithmetic), whose value lies between the given lower and upper bounds, and whose error at each arithmetic operation is independent. In Dempster (1969, 1974), a three-point arithmetic is used to approximate a discrete distribution, although there is nothing to prevent using a finer approximation except computational time consideration, and this presentation presents a synopsis of these results using the three-point approximation with instruction on how to do arithmetic on a finer mesh. In triplex arithmetic, a main value is carried. Quantile arithmetic accounts for the uncertainty within a given interval and the manner in which this uncertainty within the interval propagates. The problem of how the uncertainty is distributed is especially problematic when the uncertainty has a large support, and the bulk of the uncertainty is amassed around a single value, that is, it has a narrow dispersion and a long tail.

Developing an arithmetic begins not only with the assumption of independence, but that the variables are represented by an absolutely continuous distribution or a discrete distribution. For the three-point quantile arithmetic, consider a variable $X$, whose three-point approximation is denoted $\widehat{X}$, such that its distribution is

$$f_{\widehat{X}}(x) = \begin{cases} \alpha Y & \text{if } x = x_1, \text{ where } P(X \le x_1) = \alpha \\ 1 - 2\alpha & \text{if } x = x_2, \text{ where } P(X \le x_2) = \frac{1}{2} \\ \alpha & \text{if } x = x_3, \text{ where } P(X \le x_3) = 1 - \alpha \\ 0 & \text{otherwise} \end{cases} \tag{38}$$

with $0 \le \alpha \le \frac{1}{2}$. From the construction of $f_{\widehat{X}}$, we have $\underline{x} \le x_1 \le x_2 \le x_3 \le \bar{x}$, where $[\underline{x}, \bar{x}]$ is the support of the distribution (where it is understood that we are using the extended real line in the case in infinite support). The $x_1 \le x_2 \le x_3$ are, respectively, the $\alpha$th quantile, $\frac{1}{2}$th quantile (median), and the $(1 - \alpha)$th quantile of the absolutely continuous random variable $X$ with support $[\underline{x}, \bar{x}]$.

Let $S_\alpha$ be the space of all independent random variables of the form $\widehat{X}$, that is, whose distribution is given by Eq. (38). The parameter $\alpha$ is typically fixed *a priori*. The more concentrated the random variables in $S_\alpha$, the larger $\alpha$ may be, and the choice is based on the probabilistic interpretation that is applicable to the problem at hand. According to Dempster (1969, 1974), the choice of $\alpha = \frac{1}{20}$ is a reasonable value since 90% of the probability will then lie between $x_1$ and $x_3$. It is clear that $S_{\frac{1}{2}}$ is isomorphic to $\mathbb{R}$, and $S_0$ is the

set of triplex numbers, where the median is the main value. In what follows, it is assumed that $0 \leq \alpha < \frac{1}{2}$, and we drop the "hat" designation of the approximate random variables. Let $\circ \in \{+, -, \times, \div\}$ so that $Z = X \circ Y$ consists of a nine-point distribution, where the resultant support is handled by interval arithmetic, with discrete density

$$f_Z(z) = \begin{cases} p_i p_j & \text{if } z = x_i \circ y_j \text{ for } i, j = 1, 2, 3 \\ 0 & \text{otherwise,} \end{cases}$$

and $p_1 = \alpha$, $p_2 = 1 - 2\alpha$, $p_3 = \alpha$, where $(x_1, x_2, x_3)$ and $(y_1, y_2, y_3)$ are the defined triplets for $X$ and $Y$. We must approximate $Z$ by a triplet $(w_1, w_2, w_3)$ for it to be a member of $S_\alpha$. To do so,

1. Order the nine values $z_k = x_i \circ y_j$, $k = 1, \ldots, 9$ $(z_1 \leq \cdots \leq z_9)$, and denote their associated probabilities $q_k$.
2. Take $w_1$ to be the largest $z_k$ for which $\sum_{i=1}^{k} q_i \leq \alpha$.
3. Take $w_2$ to be the smallest $z_k$ for which $\sum_{i=1}^{k} q_i \geq \frac{1}{2}$.
4. Take $w_3$ to be the smallest $z_k$ for which $\sum_{i=1}^{k} q_i \geq 1 - \alpha$.

A real number $r$ has a distribution $R$ in $S_\alpha$ of

$$f_R(x) = \begin{cases} 1 & x = r \\ 0 & \text{otherwise,} \end{cases}$$

so that $R = (r, r, r)$. Scalar multiplication becomes

$$Z = R \times X = (r \times x_1, r \times x_2, r \times x_3).$$

Powers are also easily computed as follows:

$$\begin{aligned} Z = X^n &= (w_1, w_2, w_3), \quad \text{where} \\ w_1 &= \min\{x_1^n, x_2^n, x_3^n\}, \\ w_2 &= x_2^n, \\ w_3 &= \max\{x_1^n, x_2^n, x_3^n\}. \end{aligned}$$

Readers can find examples in Dempster (1969, page 113; and 1974, page 188). Quantile arithmetic is commutative but not associative. Real numbers, 0, and 1, are the additive and multiplicative identities. However, in general, quantile arithmetic is not even subdistributive (as is interval arithmetic). Moreover, in general, quantile arithmetic is not inclusion monotonic for all arithmetic operations. On the other hand, if $f(x_1, \ldots, x_n)$ is a rational expression, the corresponding quantile expression $F(X_1, \ldots, X_n)$ has the following important enclosure property:

$$f(x_1, \ldots, x_n) \subseteq F(X_1, \ldots, X_n).$$

The above is related to distribution arithmetic, which is the topic of Section IV.

**e. Other Interval Arithmetics.** A brief mention is made of other types of interval arithmetics and an interval-like arithmetic for the sake of completeness. These were developed to deal with dynamic modeling problems.

*Ellipsoid Arithmetic.* The ellipsoidal arithmetic of Neumaier (1993) is based on approximating enclosing affine transformations of ellipsoids that are again contained in an ellipsoid. The focus of the article is in enclosing solutions to dynamic systems models where the wrapping effect associated with interval ($n$-dimensional hyperboxes, $n \geq 2$) enclosures may severely hamper their usefulness since boxes parallel to the axes are not the optimal geometric shape to minimize bounds (Guderley and Keller, 1972). A second focus of Neumaier's article is enclosing confidence limits. Kahan (1968a) shows how to compute the "tightest" ellipsoid enclosure of the intersection of two ellipsoids. These tightest ellipsoid enclosures are the underlying basis of the approximations developed in Neumaier (1993). It is clear that computing with ellipsoids is not simple; therefore, a simple approximation is necessary if the method is to be useful. While the sum and product of ellipses are not found explicitly worked out by Neumaier, they are implicit. Enclosing the sum is straightforward. The difference, product, and quotient require approximations.

*Variable Precision Interval Arithmetic.* Variable precision interval arithmetic [Ely (1993a), Moore (1992), and more recently Revol and Rouillier (2005) and Schulte and Swartzlander (2000)] was developed to enclose solutions to problems in computational mathematics requiring more precision than afforded by the usual floating-point arithmetic (single and double precision, for example). A problem in this category is wind shear (vortex) modeling (Ely and Baker, 1993b). A specialized interval arithmetic has been developed both in software (Ely, 1993a) and in hardware (Schulte and Swartzlander, 2000). The Taylor model arithmetic (Makino and Berz, 2003, 2005) may be considered as a variable precision interval arithmetic.

## 4. *Comparison Between the Axiomatic and Extension Principle Approach to Interval Arithmetic*

The axiomatic approach to interval arithmetic considers an interval as a *number* with two components, whereas constraint interval arithmetic considers an interval as a *set*. The set point of view is the one taken by interval arithmetic that uses containment sets (Pryce and Corliss, 2006). In considering an interval as a number, interval arithmetic defines the operations axiomatically. Axiomatic interval arithmetic is simple and straightforward since it is defined through real number operations that are no less than twice, and at most four times, more complex than the corresponding real number operations. What

follows is an arithmetic that does not have additive or multiplicative inverses and is subdistributive, potentially resulting in overestimation. Exponential complexity arises in attempting to reduce overestimations.

An interval considered as a set leads to an arithmetic defined through global optimization of the united extension function of the arithmetic operations. Thus, constraint interval arithmetic requires a procedure rather than an arithmetic operation. The complexity is explicit at the onset and potentially *NP-Hard*. Nevertheless, the algebraic structure of constraint interval arithmetic not only possesses additive and multiplicative inverses, but is also distributive. It may not be easily implemented because the arithmetic is a global optimization.

Extended interval arithmetic and (nonstandard) directed interval arithmetic add axioms so that interval arithmetic can operate on the extended real number system. Generalized interval arithmetic, affine interval arithmetic, and Taylor model arithmetic deal with the dependency problem. Triplex arithmetic and quantile arithmetic carry more information in their representation so that the propagated value that results from arithmetic carries more information than just the support (end points of the interval). The specialized arithmetics (ellipsoid, variable precision, range, and rational) deal with a different representation of uncertainty. The current implementations and uses of ellipsoid and variable precision arithmetics are tailored to deal with specific problem domains.

### *C. Enclosure and Verification*

Enclosure and verification methods are approaches to problems in computational mathematics in which solutions are returned with automatically computed (computer-generated) bounds (enclosures). If the enclosure is nonempty, the goal is to verify existence and uniqueness where possible. Three different approaches to enclosure methods are presented here:

1. *Range of a function methods* compute an upper bound to a maximum and a lower bound to the minimum of a continuous function by using rounded interval arithmetic (Alefeld, 1990; Hansen, 1980, 1992; Makino and Berz, 2003, 2005; Ratschek and Rokne, 1988).
2. *Epsilon inflation methods* (Kaucher and Rump, 1982) compute an approximate solution, inflate the approximate to form an interval, and compute the range according to the above.
3. *Defect correction methods* (Böhmer *et al.*, 1984) compute an approximate inverse to the problem. If the approximate inverse composed with the given function is contractive, then iterative methods are guaranteed to converge to a solution, and mathematically correct error bounds on the solution can be computed.

The naive approach to computing the range of a rational function is to replace every algebraic operation by axiomatic interval arithmetic operations. This works in theory for continuous function with unions of smaller and smaller boxes whose diameters approach 0. However, this approach has exponential complexity. Authors have found a variety of methods to obtain the range of a function (see Hansen, 1992; Makino and Berz, 2003, 2005; Neumaier, 2004b).

The meaning of enclosure and verification in the context of interval analysis is discussed next.

***Definition 6.*** The *enclosure* of a set of real numbers (real vectors) $Y$ is a set of real numbers (real vectors) $X$ such that $Y \subseteq X$. In this case, $X$ encloses $Y$. The set $X$ is called the *enclosing set*.

Enclosure makes sense when $Y$ is an unknown, but for which bounds on its values are sought. For example, the set $Y$ could be the set of solutions to a mathematical problem. In the case of interval analysis over $\mathbb{R}$, the enclosing set $X$ is a computed interval. Typically, approximating algorithms return a real number (vector) approximation $\tilde{x}$ as the computed value of the unknown solution $y$ with no sense of the quality of the solution, that is, its error bounds. The idea of enclosure is that mathematically valid computed error bounds, $Y \subseteq X = [\underline{x}, \bar{x}]$, on the solution are provided. If the approximation to the solution is $\tilde{x} = \frac{\underline{x}+\bar{x}}{2}$, the maximal error is guaranteed to be $\text{error}_{\max} = \frac{\bar{x}-\underline{x}}{2}$. If we are dealing with functions, there are only two pertinent cases.

1. The first is the enclosure of the range of a function in a box, that is, $Y = \{f(x) \mid x \in \mathit{Domain}\} \subseteq X$, where $X$ is a box.
2. The second case is pointwise enclosure, that is, $[g(x), h(x)]$ encloses the function $f(x)$ pointwise if $g(x) \leq f(x) \leq h(x)\ \forall x \in \mathit{Domain}$. The methods to (efficiently) obtain $g(x)$ and $h(x)$ are found in Section IV.

Researchers do not give a definition to "enclosure methods," since the word *enclosure* itself seems to denote its definition. In fact, Alefeld (1990) states,

> In this paper we do not try to give a precise definition of what we mean by an enclosure method. Instead we first recall that the four basic interval operations include the range of values of rational functions. Using more appropriate tools, the range of more general functions can be included. Since all enclosures methods for solution of equations which are based on interval arithmetic tools are finally enclosures methods for the range of some function, we concentrate ourselves on methods for the inclusion of the range of function.

There is an intimate relation between enclosure and inclusion of the range of functions. However, enclosure for this study is more general than that to which Alefeld (1990) limits himself, since we deal with epsilon inflation and defect correction methods in addition to finding the range of a function.

The concept of verification for this study is restricted to the context of computed solutions to problems in continuous mathematics. Verification is defined next.

***Definition 7.*** *Verification* of solutions to a problem in continuous mathematics of $\mathbb{R}^n$ is the construction of a box $X$ that encloses the solutions of the problem in a given domain where, for $X \neq \emptyset$, at least one solution exists, and for $X = \emptyset$, no solution exists in the given domain of the problem.

Thus, verification includes the *existence* of solutions and the *computability* of enclosures. In particular, when the construction of the verified solution is carried out on a computer, the enclosures are mathematically valid enclosures whose end points are floating-point numbers. Note that even if a mathematical analysis results in $X \neq \emptyset$, it may still fail to contain a solution. For example, the computed range of a function may contain 0, but fail to have a solution to the problem $f(x) = 0$. Thus, enclosure may not mean verification. The literature often uses the term *validation* to mean what we have defined as *verification*. Methods that compute verified solutions and verify uniqueness are also called *E-methods* by Kaucher and Rump (1982), and these methods are applied to solutions of fixed-point problems $f(x) = x$. The authors develop methods to solve linear equations by *E-methods*. Many authors, in the context of verifying solutions to equations, use the word *proof* (see, for example, Section 2 of Kearfott, 1996b). While the mathematical verification of existence (and perhaps uniqueness) is a type of proof, for this monograph, the mathematical confirmation that the hypotheses of a theorem (say the Brouwer fixed-point theorem) hold is what we mean by *verification*. Nevertheless, Kearfott (1996b) states on page 3,

> A powerful aspect of interval computations is tied to the Brouwer fixed-point Theorem.

> **Theorem A** *(*Brouwer fixed point-Theorem—see any elementary text on real analysis or Neumaier (1990), page 200*). Let $D$ be a convex and compact subset of $\mathbb{R}^n$ with int$(D) \neq \emptyset$. Then every continuous mapping $G : D \rightarrow D$ has at least one fixed point $x^* \in D$, that is, a point with $x^* = G(x^*)$.*

> The Brouwer fixed-point theorem combined with interval arithmetic enables numerical verification of existence of solutions to linear and

nonlinear systems. The simplest context in which this can be explained is the one-dimensional interval Newton method.
Suppose $f : \mathbf{x} = [\underline{x}, \bar{x}] \to \mathbb{R}$ has continuous first derivative on $\mathbf{x}$, $\check{x} \in \mathbf{x}$, and $\mathbf{f}'(\mathbf{x})$ is a set that contains the range of $f'$ over $\mathbf{x}$ (such as when $f'$ is evaluated at $\mathbf{x}$ with interval arithmetic). Then the operator

$$\mathbf{N}(f; \mathbf{x}, \check{x}) = \check{x} - f(\check{x})/\mathbf{f}'(\mathbf{x}) \tag{39}$$

is termed the *univariate interval Newton method*.... Applying the Brouwer fixed-point theorem in the context of the univariate interval Newton method leads to:

**Theorem B.** *If* $\mathbf{N}(f; \mathbf{x}, \check{x}) \subset \mathbf{x}$*, then there exists a unique solution to* $f(x) = 0$ *in* $\mathbf{x}$.

Existence in Theorem B follows from *Miranda's theorem*, a corollary of the Brouwer fixed-point theorem.

Three types of verification occur in practice. These are, as mentioned, (1) enclosure of the range of a function or global optimization, (2) epsilon inflation, and (3) defect correction.

1. *Enclosure of the Range of a Function*

The enclosure of the range of a function using interval arithmetic most often assumes that a function is continuous. Thus, as long as rounded interval arithmetic is used, the resulting enclosure is verifiable to be correct (Hansen, 1992; Moore, 1966; Neumaier, 1990). Uniqueness can also be verified mathematically on a computer using methods outlined in Hansen (1992), Kearfott (1996a), Makino and Berz (2003, 2005), Neumaier (1990), or Neumaier (2004b). This article does not elaborate further on interval methods to obtain the range of a function since these methods are well represented in the literature except to point out that interval arithmetic and interval analysis have been used to compute tight constraints in artificial intelligent systems for constraint propagation. The interfaces between constraint propagation and interval analysis can be found in Lodwick (1989). It is noted that more modern methods for global optimization do use constraint propagation methods such as those found in Lodwick (1989).

2. *Epsilon Inflation*

Epsilon inflation methods are approaches for the verification of solutions to the problem $f(x) = 0$ using two steps: (1) application of a usual numerical method to solve $f(x) = 0$ to obtain an approximate solution $\hat{x}$,

and (2) inflation of $\hat{x}$ to obtain an approximate interval $\widehat{X} = [\hat{x} - \epsilon, \hat{x} + \epsilon]$, and application of interval methods using rounded interval arithmetic (for example, interval Newton's method) to obtain an enclosure. Mayer (1996, page 98) outlines how to solve problems through *E-methods* using *epsilon-inflation* techniques to solve $f(x) = 0$, where the function is assumed to be continuous over its defined domain. The idea is to solve the problem on a closed and bounded subset of its domain using the following steps:

1. Transform the problem into an equivalent fixed-point problem, $f(x) = 0 \Leftrightarrow g(x) = x$.
2. Solve the fixed point for an approximate solution $\tilde{x}$ using a known algorithm. That is, $g(\tilde{x}) \approx \tilde{x}$.
3. Identify an interval function enclosure to the fixed-point representation of the problem,

   $g(x) \in [G]\big([x]\big)\ \forall x \in [x]$, where $[x]$ is in the domain of both $g$ and $[G]$.

   For example,

$$[G]\big([x]\big) = \Big[\min_{y\in[x]} \underline{G}(y), \max_{y\in[x]} \overline{G}(y)\Big].$$

4. Verify

$$[G]\big([x]\big) \subseteq \mathit{interior}[x]$$

   by doing the following:
   (a) $[x]^0 := [\tilde{x}, \tilde{x}]$
   (b) $k = -1$
   (c) **repeat**
      i. $k := k + 1$
      ii. choose $[x]_\epsilon^k$ such that $[x]^k \subseteq \mathit{interior}([x]_\epsilon^k)$—this is the *epsilon-inflation* where $[x]_\epsilon$ is defined below.
      iii. $[x]^{k+1} := [G]([x]_\epsilon^k)$
   (d) **until** $[x]^{k+1} \subseteq \mathit{interior}([x]_\epsilon^k)$ or $k > k_{\max}$

A variety of methods can be used to pick the epsilon-inflation. In particular, Mayer (1995) uses the following:

$$[x]_\epsilon = (1 + \epsilon)[x] - \epsilon[x] + [-\eta, \eta]$$

where $\eta$ is the smallest floating-point number (machine epsilon). Another approach is as follows:

$$[y] = [\underline{y}, \bar{y}] := (1 + \epsilon)[x] - \epsilon[x],$$
$$[x]_\epsilon := \big[\mathit{pred}(\underline{y}), \mathit{succ}(\bar{y})\big]$$

where $pred(\underline{y})$ denotes the first floating-point number equal to or less than $\underline{y}$ (round down) and $succ(\bar{y})$ denotes the first floating-point number equal to or greater than $\bar{y}$ (round up). The value $\epsilon = 0.1$ has been used as an initial guess.

### 3. *Defect Correction*

*Defect correction methods* (Böhmer *et al.*, 1984) solve the fixed-point problem $f(x) = x$ by computing an approximate inverse in such a way that the approximate inverse acting on the original operator is contractive. This approach is then used in conjunction with verification (Kaucher and Rump, 1982), for example, when they are used in conjunction with epsilon-inflation and/or range enclosure outlined above. The general defect method as stated by Böhmer *et al.* (1984, page 3):

Solve

$$Fz = y, \tag{40}$$

where $F : D \subset E \to \widehat{D} \subset \widehat{E}$ is a bijective continuous, generally nonlinear operator; $E$, $\widehat{E}$ are Banach spaces. The domain and range are defined appropriately so that for every $\tilde{y} \in \widehat{D}$ there exists exactly one solution of $Fz = \tilde{y}$. The (unique) solution to Eq. (40) is denoted $z^*$. Assume that Eq. (40) cannot be solved directly, but the *defect* (also called the residual in other contexts)

$$d(\tilde{z}) := F\tilde{z} - y \tag{41}$$

may be evaluated for "approximate solutions" $\tilde{z} \in D$. Further assume that the *approximate problem*

$$\widetilde{F}z = \tilde{y} \tag{42}$$

can be readily solved for $\tilde{y} \in \widehat{D}$. That is, we can evaluate the solution operator $\widetilde{G}$ of Eq. (42). $\widehat{G} : \widehat{D} \to D$ is an *approximate inverse* of $F$ such that (in some approximate sense)

$$\widetilde{G}F\tilde{z} = \tilde{z} \quad \text{for } \tilde{z} \in D \tag{43}$$

and

$$F\widetilde{G}\tilde{y} = \tilde{y} \quad \text{for } \tilde{y} \in \widehat{D}. \tag{44}$$

Assume that an approximation $\tilde{z} \in D$ to $z^*$ is known and the defect $d(\tilde{z})$ [Eq. (41)] has been computed. There are, in general, two ways to compute another (hopefully better) approximation $\bar{z}$ to $\tilde{z}$ by solving Eq. (42):

1. Compute a change $\Delta z$ in Eq. (42) with the right-hand side being the defect, $d(\tilde{z})$, and then use $\Delta z$ as a correction for $\tilde{z}$. That is,

$$\begin{aligned} \bar{z} &:= \tilde{z} - \Delta z = \tilde{z} - \left[\widetilde{G}\big(y + d(\tilde{z})\big) - \widetilde{G}y\right] \\ \bar{z} &:= \tilde{z} - \widetilde{G}F\tilde{z} + \widetilde{G}y. \end{aligned} \tag{45}$$

   This assumes that the approximate inverse, $\widetilde{G}$, is linear, that is, $\widetilde{G}(y + d(\tilde{z})) = \widetilde{G}y + \widetilde{G}(F\tilde{z} - y) = \widetilde{G}F\tilde{z}$.
2. Use the known approximate solution $\tilde{z}$ in Eq. (42) to compute $\tilde{y}$. Now change this value by the defect to obtain $\bar{y} = \tilde{y} - d(\tilde{z})$. Use the approximate inverse and solve using $\bar{y}$. That is,

$$\begin{aligned} \bar{y} &:= \tilde{y} - d(\tilde{z}) = \tilde{y} - (F\tilde{z} - y) \\ &= \tilde{y} - F\widetilde{G}\tilde{y} + y, \end{aligned}$$

   since $\tilde{y} = F\tilde{z}$, so that from Eq. (43) $\widetilde{G}\tilde{y} = \widetilde{G}F\tilde{z} = \tilde{z}$, that is, $\tilde{z} = F\widetilde{G}\tilde{y}$. Now, the new approximation to $\tilde{z}$ becomes

$$\bar{z} = \widetilde{G}\bar{y} = \widetilde{G}\big[(\widetilde{F} - F)\tilde{z} + y\big], \tag{46}$$

   where again, we must assume that the inverse operator $\widetilde{G}$ is linear.

The success of the defect correction, steps 1 [Eq. (45)], or 2 [Eq. (46)], depends on the *contractivity* of the operators

$$(I - \widetilde{G}F) : D \to D,$$

or

$$(I - F\widetilde{G}) : \widehat{D} \to \widehat{D},$$

respectively, since Eq. (45) implies

$$\bar{z} - z^* = (I - \widetilde{G}F)\tilde{z} - (I - \widetilde{G}F)z^*,$$

whereas Eq. (46) implies

$$\bar{y} - y^* = (I - F\widetilde{G})\tilde{y} - (I - F\widetilde{G})y^*.$$

The associated iterative algorithm is (Stetter, 1978):

**DEFECT CORRECTION 1** [Eq. (45)]

$$z_{k+1} = z_k - \widetilde{G}Fz_k + \widetilde{G}y \tag{47}$$

**DEFECT CORRECTION 2** [Eq. (46)]

$$\begin{aligned} y_{k+1} &= y_k - F\widetilde{G}y_k + y \\ z_k &= \widetilde{G}y_k \\ z_{k+1} &= \widetilde{G}\big[(\widetilde{F} - F)z_k + y\big]. \end{aligned} \tag{48}$$

### *D. Algorithms and Software*

The practical issues of using intervals to solve mathematical problems must include the following: (1) how to handle dependencies (repeated occurrences of the same variable), (2) how to obtain enclosures when carried out on a particular computer in such a way that they are mathematically correct, and (3) efficient algorithms and computer structures in software systems. Methods have been presented that deal with dependencies, enclosures, verifiability, and computability. For example, iterative methods are particularly well suited to deal with computability problems. Approximate global optimization techniques, even NP-Hard, can obtain good approximations using constraint propagation and other interval techniques (Neumaier, 2004b). As mentioned in Corliss (2004, pages 35–36), quality production packages that incorporate some or all the above issues include:

1. INTLAB, which is a Matlab interval analysis system that can be downloaded from http://www.ti3.tu-harburg.de/~rump/intlab/index.html
2. PROFIL/BIAS (Programmer's Runtime Optimized Fast Interval Library/Basic Interval Arithmetic Subroutines) for Matlab from TU Hamburg-Harburg (http://www.ti3.tu-harburg.de/Software/PROFILEnglisch.html).
3. Fortran 95 and C++ from Sun Microsystems (wwws.sun.com/software/sundev/suncc/index.html). Moreover, interval arithmetic as a native type has been on SPARC/Solaris since about 2000. It is free (Sun Studio 11 release) and can be obtained at http://developers.sun.com/prodtech/cc/index.jsp. Documentation can be downloaded from http://docs.sun.com/app/docs/doc/819-3695. The C++ documentation can be downloaded from http://docs.sun.com/app/docs/doc/819-3696.
4. C-XSC, Pascal-XSC packages from TU Karlruhe at www.uni-karlsruhe.de/~iam/html/language/xsc-sprachen.html
5. FI-LIB, the authors, Hofschuster and Krämer, state that FI_LIB is a "fast interval library (version 1.2) in ANSI-C…" whose main features of the library, called fi_lib (fast interval library) are:
   (a) Fast table look-up algorithms are used for the basic functions, such arctan, exp, or log.
   (b) All elementary function routines are supplied with reliable relative error bounds of high quality. The error estimates cover rounding errors, introduced by not exactly representable constants, as well as approximation errors (best approximations with reliable error bounds).
   (c) All error estimates are reliable worst-case estimates, which have been derived using interval methods.

(d) We only insist on faithful computer arithmetic. The routines do not manipulate the rounding mode of basic operations (setting the rounding mode may be rather expensive).
(e) No higher-precision internal data format is used. All computations are done using the IEEE double format (64-bit).
(f) A C++ interface for easier use is also supplied with the library.
(g) For good portability, all programs are written in ANSI-C.

6. FILIB++ (http://www.math.uni-wuppertal.de/wrswt/software/filib.html), FILIB++ Interval Library whose authors are Lerch, Tischler, Wolff von Gudenberg, Hofschuster, and Krämer. They say that, "filib++ is an extension of the interval library filib. The most important aim of the latter was the fast computation of guaranteed bounds for interval versions of a comprehensive set of elementary function. filib++ extends this library in two aspects. First, it adds a second mode, the "extended" mode, that extends the exception-free computation mode using special values to represent infinities and NotaNumber known from the IEEE floating-point standard 754 to intervals. In this mode so-called containment sets are computed to enclose the topological closure of a range of a function defined over an interval. Second, state-of-the-art design uses templates and trait classes in order to get an efficient, easily extendable and portable library, fully according to the C++ standard."
7. Maple (see http://www.maplesoft.com) and Mathematica (see http://www.wri.com) support interval structures. Wagon (Bornemann *et al.*, 2004) and colleagues solved a suite of challenge problems using Mathematica's interval subroutines.
8. COSY, from Michigan State University, is a language for verified global optimization and ordinary differential equation solving, based on intervals as well as Taylor models with remainder bounds (see http://www.beamtheory.nscl.msu.edu/cosy).
9. The interval analysis website (http://www.cs.utep.edu/interval-comp/intlang.html), has a list of languages that support interval data types.
10. Global optimization systems that use interval analysis are found in the software developed by the COCONUT project. "COCONUT is an IST Project funded by the European Union. Its goal is to integrate the currently available techniques from mathematical programming, constraint programming, and interval analysis into a single discipline, to get algorithms for global optimization and continuous constraint satisfaction problems that outperform the current generation of algorithms based on using only techniques from one or two of the traditions." (See http://www.mat.univie.ac.at/users/neum/public_html/glopt/coconut/.) Another global optimization solver, GLOBSOV, can be downloaded from http://interval.louisiana.edu/GlobSol/download_GlobSol.html.

11. UniCalc (see http://archives.math.utk.edu/software/msdos/miscellaneous/unicalc/) is another system that uses intervals. Its website states that, "UniCalc is a solver for mathematical problems using novel mathematical tools for calculations. This approach makes it possible to solve principally new classes of problems and to obtain new results for problems solved with standard methods of calculations. UniCalc is intended to solve direct and inverse problems represented by systems of algebraic and differential equations, inequalities, and logic expressions. The system may be overdetermined or underdetermined, and the system's parameters can be given imprecisely. UniCalc allows calculations to be performed both with integers and real variables that may be mixed in the system. As a result of calculations, a set of intervals containing all real solutions of the system is found."

## III. Fuzzy Set Theory

Fuzzy set and possibility theory were defined and developed by Zadeh beginning in 1968 with subsequent articles in 1968 and 1975. As is now well known, the idea was to mathematize and develop analytical tools to solve problems whose uncertainty was more ample in scope than probability theory. Classical mathematical sets—for example, a set $A$—have the property that either an element $x \in A$, or $x \notin A$, but not both. There are no other possibilities for classical sets, which are also called *crisp* sets. An interval is a classical set. Zadeh's idea was to relax this "all-or-nothing" membership in a set to allow for grades of belonging to a set. When grades of belonging are used, a fuzzy set ensues. To each fuzzy set, $\widetilde{A}$, Zadeh associated a real-valued function $\mu_{\widetilde{A}}(x)$, called a *membership function*, for all $x$ in the domain of interest, the universe $\Omega$, whose range is in the interval $[0, 1]$ that describes, quantifies the degree to which $x$ belongs to $\widetilde{A}$. For example, if $\widetilde{A}$ is the fuzzy set "middle-aged person," then a 15-year-old has a membership value of zero whereas a 35-year-old might have a membership value of 1, and a 40-year-old might have a membership value of $1/2$. That is, a fuzzy set is a set for which membership in the set is defined by its membership function $\mu_{\widetilde{A}}(x) : \Omega \to [0, 1]$ where a value of 0 means that an element does not belong to the set $\widetilde{A}$ with certainty and a value of 1 means that the element belongs to the set $\widetilde{A}$ with certainty. Intermediate values indicate the degree to which an element belongs to the set. Using this definition, a classical (so-called crisp) set $A$ is a set whose membership function has a range that is binary, that is, $\mu_A(x) : \Omega \to \{0, 1\}$, where $\mu_A(x) = 0$ means that $x \notin A$, and $\mu_A(x) = 1$ means $x \in A$. This membership function for a crisp set $A$ is, of course, the characteristic function. So a fuzzy set can be thought of as being one that has

a generalized characteristic function that admits values in [0, 1] and not just two values {0, 1} and is uniquely defined by its membership function. Another way of looking at a fuzzy set is as a set in $\mathbb{R}^2$ as follows.

***Definition 8.*** A fuzzy set $\widetilde{A}$, as a crisp set in $\mathbb{R}^2$, is the set of ordered pairs

$$\widetilde{A} = \{(x, \mu_{\widetilde{A}}(x))\} \subseteq \{(-\infty, \infty) \times [0, 1]\}. \tag{49}$$

Some of the earliest people to recognize the relationship between interval analysis and fuzzy set theory were Nguyen (1978), implicitly, Dubois and Prade (1980, 1981), and Kaufmann and Gupta (1985), explicitly. In particular, Dubois and Prade (1987a, 1991), Dubois *et al.* (2000b), Fortin *et al.* (2006) deal specifically with interval analysis and its relationship with fuzzy set theory. In Dubois *et al.* (2000b), it is shown that,

> ... set-inclusive monotonicity, as given by R.E. Moore (see Moore, 1966, 1979), holds for fuzzy quantities. That is, for fuzzy sets $\widetilde{A}$ and $\widetilde{B}$,
>
> $$\widetilde{A} \subseteq \widetilde{B} \quad \Longrightarrow \quad f(\widetilde{A}) \subseteq f(\widetilde{B}).$$
>
> This crucial result just reminds us that when the operands become more imprecise, the precision of the result cannot but diminish. Due to its close relationship to interval analysis, the calculus of fuzzy quantities is clearly pessimistic about precision, since $f(A_1, A_2)$ is the largest fuzzy set in the sense of fuzzy set inclusion, that is,
>
> $$\widetilde{A} \subseteq \widetilde{B} \quad \Longleftrightarrow \quad \mu_A(x) \leq \mu_B(x), \quad \forall x.$$

Much has been written about fuzzy sets that can be found in standard textbooks (Klir and Yuan, 1995); this material is not repeated here. We present only the ideas that are pertinent to the areas in the interfaces between interval and fuzzy analysis of interest. Given that the primary interest is in the relationships between real-valued interval and fuzzy analysis, we restrict our fuzzy sets to a real-valued universe, $\Omega \subseteq \mathbb{R}$, whose membership functions, *fuzzy numbers*, or *fuzzy intervals*, are defined below.

***Definition 9.*** A *modal value* of a membership function is a domain value at which the membership function is one. A fuzzy set with at least one modal value is called *normal*. The *support* of a membership function is the closure of $\{x \mid \mu_{\widetilde{A}}(x) > 0\}$.

***Definition 10*** (Fortin *et al.*, 2006). A *fuzzy interval*, $M$, defined by its membership function $\mu_M(\cdot)$, is a fuzzy subset of the real line such that, if $x, y, z \in \mathbb{R}$, $z \in [x, y]$, then

$$\mu_M(z) \geq \min\{\mu_M(x), \mu_M(y)\}.$$

Like a fuzzy set, a fuzzy interval $M$ is said to be *normal* if $\exists x \in \mathbb{R}$, such that $\mu_M(x) = 1$. The set $\{x \mid \mu_M(x) = 1\}$ is called the *core* of the fuzzy interval.

For all that follows, fuzzy intervals will be assumed to be normal fuzzy intervals with upper semi-continuous membership functions. This means that the $\alpha$-*cut* of a fuzzy interval,

$$M_\alpha = \big\{x \mid \mu_M(x) \geq \alpha > 0\big\},$$

is a closed interval. Let $M_1 = \{x \mid \mu_M(x) = 1\} = [m_1^-, m_1^+]$ be the core of a fuzzy interval $M$, where the support is $M_0 = \{x \mid \mu_M(x) > 0\} = (m_0^-, m_0^+)$. For a fuzzy interval, $M$, $\mu_M(x)$, for $x \in (-\infty, m_1^-]$ is nondecreasing, and $\mu_M(x)$, for $x \in [m_1^-, \infty)$ is nonincreasing. The definition of fuzzy number is given next.

***Definition 11.*** A *fuzzy number* is a fuzzy interval with a unique modal value, that is, the core is a singleton.

The fact that we have closed intervals at each $\alpha$-*cut* means that fuzzy arithmetic can be defined by interval arithmetic on each $\alpha$-*cut*, as will be seen. Unbounded intervals can be handled by extended interval arithmetic. In fact, when dealing with fuzzy intervals, the operations and analysis can be considered as interval operations and analysis on $\alpha$-*cuts*. However, a different and more recent approach is possible. Instead of considering a fuzzy number as a specialized fuzzy set over the set of real numbers, $\mathbb{R}$, Dubois and Prade (2005) and Fortin *et al.* (2006) revise the theory of fuzzy numbers so that a (real-valued) fuzzy number is to a (real-valued) interval what a fuzzy set is to a (classical) set. To this end, what are called *gradual numbers* were created (Dubois and Prade, 2005; Fortin *et al.*, 2006).

***Definition 12*** (Fortin *et al.*, 2006). A *gradual number* $\tilde{r}$ is defined by an assignment $A_{\tilde{r}}$ from $(0, 1]$ to $\mathbb{R}$.

The interest of this article is on special assignments that are associated with fuzzy intervals. The idea will be to define a gradual number associated with a fuzzy interval by the inverse of two functions; one is the inverse of the membership function restricted to $(-\infty, m_1^-]$, that is the inverse of

$$\mu_{\widetilde{A}}^-(x) = \mu_{\widetilde{A}}(x), \quad x \in \big(-\infty, m_1^-\big],$$

where $[m_1^-, m_1^+]$ is the core as before, which is nonempty. The second function is the inverse of the membership function restricted to $[m_1^+, \infty)$, that

is, the inverse of

$$\mu^{+}_{\widetilde{A}}(x) = \mu_{\widetilde{A}}(x), \quad x \in \left[m_1^{+}, \infty\right).$$

These inverses,

$$\left(\mu^{-}_{\widetilde{A}}\right)^{-1}(\alpha) : (0, 1] \to \mathbb{R} \tag{50}$$

and

$$\left(\mu^{+}_{\widetilde{A}}\right)^{-1}(\alpha) : (0, 1] \to \mathbb{R}, \tag{51}$$

define the gradual numbers in the context of real fuzzy intervals, which is our interest. Thus, for a fuzzy interval, $\widetilde{A}$, the functions $(\mu^{-}_{\widetilde{A}})^{-1}(\alpha)$ [Eq. (50)] and $(\mu^{+}_{\widetilde{A}})^{-1}(\alpha)$ [Eq. (51)], are special cases of this definition, and we concentrate on fuzzy sets that describe fuzzy intervals but with some restrictions that are specified next.

***Definition 13*** (Fortin *et al.*, 2006). Using the notion of gradual number, we can describe a fuzzy interval $M$ by an ordered pair of gradual numbers $(\tilde{m}^{-}, \tilde{m}^{+})$, where $\tilde{m}^{-}$ is called the fuzzy lower bound, or *left profile*, and $\tilde{m}^{+}$ is called the fuzzy upper bound, or *right profile*.

To ensure that the left and right profiles adhere to what has been defined as a *fuzzy* interval (as opposed to an interval), several properties of $\tilde{m}^{-}$ and $\tilde{m}^{+}$ must hold. In particular (Fortin *et al.*, 2006):

1. The domains of the assignment functions, $A_{\tilde{m}^{-}}$ and $A_{\tilde{m}^{+}}$, must be in (0, 1].
2. $A_{\tilde{m}^{-}}$ must be increasing and $A_{\tilde{m}^{+}}$ must be decreasing.
3. $\tilde{m}^{-}$ and $\tilde{m}^{+}$ must be such that $A_{\tilde{m}^{-}} \leq A_{\tilde{m}^{+}}$.

***Remark.*** Fuzzy intervals with properties 1–3 above possess well-defined inverses that are functions. Note that an interval, $[a, b]$, has constant assignments, that is, $A_{\tilde{m}^{-}}(\alpha) = a$ and $A_{\tilde{m}^{+}}(\alpha) = b, 0 < \alpha \leq 1$. Since it is constant, it contains no fuzziness and is simply an interval, not a fuzzy interval. What in the literature is called a trapezoidal fuzzy number is indeed a fuzzy interval (the left profile is strictly linearly increasing and the right profile is strictly linearly decreasing)—that is, an interval has no fuzziness in the left/right profiles (they are horizontal line segments). A nonhorizontal assignment function indicates fuzziness present.

The properties that are associated with gradual numbers applied to fuzzy intervals that are relevant to this study are as follows (Fortin *et al.*, 2006):

1. Gradual numbers display nonsharpness when their assignments are not constant.
2. Gradual numbers are what Dubois and Prade (2005) call *fuzzy elements*.
3. Gradual numbers (an assignment) do not account for incomplete information. Pairs of gradual numbers may be used to account for incomplete information. In fact, a fuzzy interval (which does account for incomplete information) is defined as a pair of gradual numbers.
4. A gradual number is not a fuzzy set.
5. Selecting a gradual number in a fuzzy interval is the selection of an element per $\alpha$-*cut*. That is, it is an assignment that takes one value for each $\alpha$-*cut*.

### A. *Possibility and Necessity Distribution*

Two types of measures and distributions are associated with fuzzy sets, measures, and membership functions—possibility and necessity measures and distributions. Whereas fuzzy measures quantify the uncertainty of gradualness, possibility and necessity measures are ways to quantify the uncertainty of lack of information. Books and articles are available that develop possibility theory (see Dubois and Prade, 1980; Dubois and Prade, 2000a; Klir and Yuan, 1995; Wang and Klir, 1992). What is of interest here for fuzzy and possibilistic mathematical analysis, is called *quantitative* possibility theory (Dubois *et al.*, 2000b). General possibility theory may be derived at least in any one of the following ways:

1. Through *normalized* fuzzy sets (see Zadeh, 1975).
2. Axiomatically, from *fuzzy measures* $g$ that satisfy (Dubois and Prade, 1988; Klir and Yuan, 1995)

$$g(A \cup B) = \max\{g(A), g(B)\}.$$

3. Through the belief functions of Dempster–Shafer theory, whose focal elements are normalized and nested (Klir and Yuan, 1995).
4. By construction, through nested sets with *normalization*; for example, nested $\alpha$-*level* sets (Jamison and Lodwick, 2002).

The third and fourth approaches are of special interest since they lead directly into a quantitative possibility theory, although the first points the way and lays the foundation. It will be assumed that the possibilistic/necessity measures and distributions that are used herein are constructed according to the third or fourth approaches.

### *B. Semantics of Fuzzy Sets and Possibility and Necessity Distributions*

Confusion often exists about the differences between fuzzy and possibilistic analysis. Fuzzy and possibilistic entities have a different development from the first principles as noted previously. Moreover, they have different meanings, or semantics. Fuzzy and possibility uncertainty model different entities. Fuzzy entities, as is well known, are sets with nonsharp boundaries in which there is a transition between elements that belong and elements that do not belong to the set. Possibilistic entities are entities that exist, but the evidence associated with whether a particular element is the entity or not is incomplete or hard to obtain. Quantitative possibility distributions constructed from first principles require nested sets (Jamison and Lodwick, 2002) and normalization. Possibility distributions are normalized since their semantics are tied to existent entities. Normalization is not required of fuzzy membership functions. Thus, *not all fuzzy sets can give rise to possibility distributions*. That is, even though Zadeh's original development of possibility theory was derived from fuzzy sets, possibility theory is different from fuzzy set theory.

Possibilistic distributions (of fuzzy numbers) encapsulate the best estimate of the possible values of an entity given the available information. Fuzzy membership function values (of fuzzy numbers) describe the degree to which an entity is that value. Note that if the possibility distribution at $x$ is 1, this signifies that the best evidence available indicates $x$ is the entity that the distribution describes. On the other hand, if the fuzzy membership function value at $x$ is 1, $x$ is *certainly* the value of the entity that the fuzzy set describes. Thus, the nature of mathematical analysis, in the presence of fuzzy and possibilistic uncertainties, is quite different semantically.

The most general form of possibility theory (the first and second approaches to possibility theory listed above) establishes an order among variables with respect to the potential of their being an entity. The *magnitudes* associated with this ordering have no significance other than an indication of order. Thus, if $possibility_A(x) = 0.75$ and $possibility_A(y) = 0.25$, all that can be said is that the evidence is stronger that $x$ is the entity $A$ than $y$. One *cannot* conclude that $x$ is three times more likely to be $A$ than $y$ is. This means that for mathematical analysis, if the possibility distributions were constructed using the most general assumptions, comparisons among several distributions are restricted (to merely order). For the most general possibility theory, setting the possibility level to be greater than or equal to a certain fixed value $\alpha$, $0 \leq \alpha \leq 1$, does not have the same meaning as setting a probability to be at least $\alpha$. In the former case, the $\alpha$ has no inherent meaning (other than if one has a $\beta > \alpha$, one prefers the decision that generated $\beta$ to that which

generated $\alpha$), whereas in the latter, the *value* of $\alpha$ is meaningful. The third and fourth derivations of possibility theory lead to quantitative possibility theory.

An alternative approach to possibility theory is as a system of lower and upper distributions bounding a given, yet unknown probability. That is, given a measurable set $A$, $Nec(A) \leq prob(A) \leq Pos(A)$ bounds the unknown probability of the event $A$ so that $Pos(A) \leq \alpha$ guarantees that $prob(A) \leq \alpha$. If the possibilistic entities are constructed from this perspective, then their *$\alpha$-levels* are numerically meaningful beyond simply being an ordering. This is the method developed in Jamison and Lodwick (2002).

### C. *Fuzzy Extension Principles*

Fuzzy extension principles show how to transform real-valued functions into functions of fuzzy sets. The meaning of the associated arithmetic depends directly on the extension principle in force, since arithmetic operations are (continuous) functions over the reals, assuming division by 0 is not allowed, and over the extended reals (Hansen, 1975) when division by 0 is allowed. The fuzzy arithmetic coming from Zadeh's extension principle (Zadeh, 1965), and its relationship to interval analysis, has an extensive development (see, for example, Kaufmann and Gupta, 1985). Moreover, there is an intimate interrelationship between the extension principle being used and the analysis that ensues. For example, in optimization, the manner in which union and intersection are extended via *t-norms* and *t-conorms* will determine the constraint sets so that the way trade-offs among decisions is made is captured (see Kaymak and Sousa, 2003).

The extension principle in the context of fuzzy set theory was first proposed, developed, and defined in Zadeh (1965, 1975).

***Definition 14*** (Extension Principle of Zadeh, 1965, 1975). Given a real-valued function, $f : X \to Y$, the function over fuzzy sets $F : S(X) \to S(Y)$, where $S(X)$ [respectively, $S(Y)$] is the set of all fuzzy sets of $X$ (respectively, $Y$) is given by

$$\mu_{F(A)}(y) = \sup\{\mu_A(x) \mid y = f(x)\} \tag{52}$$

for all fuzzy subsets $A$ of $S(X)$. In particular, if $(X_1, \ldots, X_n)$ is a vector of fuzzy intervals, and $f(x_1, \ldots, x_n)$ a real-valued function, then

$$\mu_{F(X_1,\ldots,X_n)}(y) = \sup_{(x_1,\ldots,x_n)\in(X_1,\ldots,X_n)} \min_i \{\mu_{X_i}(x_i) \mid i = 1, \ldots, n \text{ and } y = f(x_1, \ldots, x_n)\}. \tag{53}$$

This definition [Eq. (53)] of the extension principle has led to axiomatic fuzzy arithmetic (corresponding to axiomatic interval arithmetic). Moreover, it is one of the main mechanisms in the literature used for fuzzy interval analysis. Various researchers have dealt with the issue of the extension principle and have amplified its applicability. In his 1978 paper, Nguyen pointed out that a fuzzy set needs to be defined to be what Dubois and Prade later called a fuzzy interval (Dubois and Prade, 2005; Fortin *et al.*, 2006) in order that

$$\left[f(A, B)\right]_\alpha = f(A_\alpha, B_\alpha),$$

where the function, $f$, is assumed to be continuous. In particular, $A_\alpha$ and $B_\alpha$ need to be compact (i.e., closed/bounded intervals) for each $\alpha$-*cut*. Thus, Nguyen defined a fuzzy number as one whose membership function is upper semi-continuous and for which the closure of the support is compact. In this case, the $\alpha$-*cuts* generated are closed and bounded (compact) sets, that is, real-valued intervals. This is a well-known result in real analysis; that is, when $f$ is continuous, the decomposition by $\alpha$-*cuts* can be used to compute $f(X_1, \ldots, X_n)$ through interval analysis by Nguyen (1978) as

$$\left[f(X_1, \ldots, X_n)\right]_\alpha = f\left([X_1]_\alpha, \ldots, [X_n]_\alpha\right).$$

It should be noted that, considering a fuzzy interval as a particular pair of gradual numbers (left and right profiles), the extension principle may be accomplished without using $\alpha$-*cuts*.

Yager (1986) pointed out that, by looking at functions as graphs (in the Euclidean plane), the extension principle could be extended to include all graphs, thus allowing for analysis of what he calls "nondeterministic" mappings, that is, graphs that are not functions. Now, "nondeterminism," as used by Yager, can be considered as point-to-set mappings. Thus, Yager implicitly restores the extension principle to a more general setting of point-to-set mappings.

Ramik (1986) points out that we can restore Zadeh's extension principle to its most general setting of set-to-set mappings explicitly. In fact, a fuzzy mapping is indeed a set-to-set mapping. He defines the image of a fuzzy set-to-set mapping as being the set of $\alpha$'s generated by the function on the $\alpha$-*cuts* of domain.

Finally, Lin's paper (Lin, 2005) is concerned with determining the function space in which the fuzzy set generated by the extension principle "lives"; that is, the extension principle generates the resultant membership function in the range space. Suppose one is interested in stable controls; one way to extend is to generate resultant (range space) membership functions that are continuous. The definition of continuous function states that small perturbations in the input (i.e., domain) cause small perturbations in the output (i.e., range), which

is one way to view the definition of stability. Lin points out conditions that are necessary in order that range membership functions have some desired characteristics (such as continuity or smoothness).

These extension principles express how to define functions over fuzzy sets, so that the resulting range has various properties of interest and defining what may be done in the space of where the extension sends the fuzzy set through the function as dictated by the extension principle itself.

1. *Fuzzy Arithmetic*

Fuzzy arithmetic was, like interval arithmetic, derived from the extension principle of Zadeh (1965). Nahmias (1978) defined fuzzy arithmetic via the fuzzy convolution as follows:

1. **Addition:** $\mu_{Z=X+Y}(z) = \sup \min_x \{\mu_X(x), \mu_Y(z-x)\}$, where $z = x + y$.
2. **Subtraction:** $\mu_{Z=X-Y}(z) = \sup \min_x \{\mu_X(x), \mu_Y(x-z)\}$, where $z = x - y$.
3. **Multiplication:** $\mu_{Z=X\times Y}(z) = \sup \min_x \{\mu_X(x), \mu_Y(z/x)\}$, where $z = x \times y$.
4. **Division:** $\mu_{Z=X\div Y}(z) = \sup \min_x \{\mu_X(x), \mu_Y(x/z)\}$, where $z = x \div y$.

The arithmetic of fuzzy entities was originally conceived by the above definition. When the extension principle of Zadeh (1965) was applied to 1–4 above, assuming that the fuzzy entities involved were noninteractive (independent), what has come to be known as fuzzy arithmetic followed. At least by 1975, deriving fuzzy arithmetic as interval arithmetic on $\alpha$-*cuts* had been described from the work of Negoita and Ralescu (1975). Also, by 1975, Zadeh (1975) had already used the extension principle to define fuzzy arithmetic. Combining these two research streams (the extension principle and arithmetic on $\alpha$-*cuts* to define arithmetic on fuzzy numbers), fuzzy arithmetic developed into interval arithmetic on $\alpha$-*cuts* (Dubois and Prade, 1977, 1978, 1980, 1981). Moore is explicitly mentioned in Dubois and Prade (1980). Much of what occurred to interval arithmetic occurred to fuzzy arithmetic—its roots in the extension principle were eliminated when given noninteraction (independence), and the axioms for arithmetic using Nguyen (1978) and requiring membership functions to be lower/upper semi-continuous.

***Definition 15.*** A function $f : \mathbb{R} \to \mathbb{R}$ is *upper semi-continuous* at $x_0$ if

$$\limsup_{x \to x_0} f(x) \leq f(x_0).$$

Fuzzy arithmetic using interval arithmetic on $\alpha$-*cuts* saw its full development in 1985 (Kaufmann and Gupta, 1985). In fact, 1985 was the year

that R.E. Moore gave a plenary talk at the first International Fuzzy System Association Congress in Mallorca, Spain.

**a. Axiomatic Fuzzy Arithmetic.** The fuzzy arithmetic developed by Kaufmann and Gupta (1985) is taken as the standard approach, while a more recent approach is found in Hanss (2005). What is needed is the fact that a fuzzy interval is uniquely determined by its $\alpha$-*cuts*, $\widetilde{A} = \bigcup_{\alpha\in(0,1]}[\mu^{-}_{\widetilde{A}}(\alpha), \mu^{+}_{\widetilde{A}}(\alpha)]$, where $\mu^{-}_{\widetilde{A}}(\alpha)$ and $\mu^{+}_{\widetilde{A}}(\alpha)$ are the left/right end points of the $\alpha$-*cuts* of the fuzzy set $\widetilde{A}$. In particular, for fuzzy intervals we have:

$$\widetilde{A} + \widetilde{B} = \bigcup_{\alpha\in(0,1]} \left\{\left[\mu^{-}_{\widetilde{A}}(\alpha), \mu^{+}_{\widetilde{A}}(\alpha)\right] + \left[\mu^{-}_{\widetilde{B}}(\alpha), \mu^{+}_{\widetilde{B}}(\alpha)\right]\right\}, \tag{54}$$

$$\widetilde{A} - \widetilde{B} = \bigcup_{\alpha\in(0,1]} \left\{\left[\mu^{-}_{\widetilde{A}}(\alpha), \mu^{+}_{\widetilde{A}}(\alpha)\right] - \left[\mu^{-}_{\widetilde{B}}(\alpha), \mu^{+}_{\widetilde{B}}(\alpha)\right]\right\}, \tag{55}$$

$$\widetilde{A} \times \widetilde{B} = \bigcup_{\alpha\in(0,1]} \left\{\left[\mu^{-}_{\widetilde{A}}(\alpha), \mu^{+}_{\widetilde{A}}(\alpha)\right] \times \left[\mu^{-}_{\widetilde{B}}(\alpha), \mu^{+}_{\widetilde{B}}(\alpha)\right]\right\}, \tag{56}$$

$$\widetilde{A} \div \widetilde{B} = \bigcup_{\alpha\in(0,1]} \left\{\left[\mu^{-}_{\widetilde{A}}(\alpha), \mu^{+}_{\widetilde{A}}(\alpha)\right] \div \left[\mu^{-}_{\widetilde{B}}(\alpha), \mu^{+}_{\widetilde{B}}(\alpha)\right]\right\}. \tag{57}$$

For fuzzy sets whose membership functions are semi-continuous,

$$(\widetilde{A} \circ \widetilde{B})_\alpha = (\widetilde{A})_\alpha \circ (\widetilde{B})_\alpha, \quad \circ \in \{+, -, \times, \div\}.$$

Computer implementation of Eqs. (54)–(57) can be found in Anile *et al.* (1995). This program uses INTLAB (another downloadable system that has interval data types and runs in conjunction with MATLAB) to handle the fuzzy arithmetic on $\alpha$-*cuts*.

**b. Case-Based Fuzzy Arithmetic.** Klir (1997) notices, as Moore before him, that if Eqs. (54)–(57) were used, overestimations will occur. Moreover, when this approach is used, $\widetilde{A} - \widetilde{A} \neq 0$ and $\widetilde{A} \div \widetilde{A} \neq 1$. Klir's idea for fuzzy arithmetic, with requisite constraints, is to do fuzzy arithmetic using constraints dictated by the context of the problem. That is, Klir defines exceptions to obtain $\widetilde{A} - \widetilde{A} = 0$ and $\widetilde{A} \div \widetilde{A} = 1$.

**c. Constraint Fuzzy Arithmetic.** Klir's (1997) approach to fuzzy arithmetic requires an *a priori* knowledge (through cases) of which variables are identical. Constraint fuzzy arithmetic (Lodwick, 1999) carries this information in the parameters, that is, it performs Eqs. (54)–(57) using a parameter, $\lambda_x$, that identifies the variable. The resulting fuzzy arithmetic derived from constraint interval arithmetic on $\alpha$-*cuts* is essentially fuzzy arithmetic with requisite constraints of Klir without cases.

**d. Fuzzy Arithmetic Using Gradual Numbers (Fortin *et al.*, 2006).** The implementation of Fortin *et al.* (2006) as a way to perform fuzzy arithmetic uses Eqs. (50) and (51) in the following way:

$$\widetilde{A} \circ \widetilde{B} = \begin{cases} (\mu^{-}_{\widetilde{A}\circ\widetilde{B}})^{-1}(\alpha) \\ \quad = \min\{(\mu^{-}_{\widetilde{A}})^{-1}(\alpha) \circ (\mu^{-}_{\widetilde{B}})^{-1}(\alpha), (\mu^{-}_{\widetilde{A}})^{-1}(\alpha) \circ (\mu^{+}_{\widetilde{B}})^{-1}(\alpha), \\ \qquad (\mu^{+}_{\widetilde{A}})^{-1}(\alpha) + (\mu^{-}_{\widetilde{B}})^{-1}(\alpha)(\mu^{+}_{\widetilde{A}})^{-1}(\alpha) \circ (\mu^{+}_{\widetilde{B}})^{-1}(\alpha)\} \\ (\mu^{+}_{\widetilde{A}\circ\widetilde{B}})^{-1}(\alpha) \\ \quad = \max\{(\mu^{-}_{\widetilde{A}})^{-1}(\alpha) \circ (\mu^{-}_{\widetilde{B}})^{-1}(\alpha), (\mu^{-}_{\widetilde{A}})^{-1}(\alpha) \circ (\mu^{+}_{\widetilde{B}})^{-1}(\alpha), \\ \qquad (\mu^{+}_{\widetilde{A}})^{-1}(\alpha) + (\mu^{-}_{\widetilde{B}})^{-1}(\alpha)(\mu^{+}_{\widetilde{A}})^{-1}(\alpha) \circ (\mu^{+}_{\widetilde{B}})^{-1}(\alpha)\} \end{cases}$$

for $\circ \in \{+, -, \times, \div\}$. Fuzzy arithmetic, using gradual numbers, can handle dependencies in many cases.

**e. Propagation.** The application of the extension principle to the evaluation of functions can be considered as propagation of uncertainty. This matter has been explicitly studied by Lodwick and Jamison (2003b), and later by Baudrit *et al.* (2005). The next section considers this topic under distribution arithmetic, where the general theory for propagation of more than one of the uncertainties of interest is the last subsection of the next section, covering in theoretical terms what Lodwick and Jamison (2003b) and Baudrit *et al.* (2005) introduce.

### *D. Enclosure and Verification*

Enclosures require the acquisition of lower and upper bounds. For fuzzy, possibilistic, and probabilistic uncertainty, lower and upper bounds are lower and upper distribution functions. Since there are distinct methods for this, and the complexity of the problem is significant, we discuss enclosure and verification for fuzzy, possibilistic, and probabilistic uncertainty as a separate section next. At the heart of enclosure and verification methods is distribution arithmetic, as interval arithmetic is at the heart of enclosure and verification in interval analysis.

## IV. Analysis with Distributions

This section extends the results of the previous section to general distributions, not just intervals and fuzzy sets, for the purpose of doing mathematical analysis on entities associated with the types of uncertainties that are of interest to this monograph. Mathematical analysis that includes a mixture

of uncertainty types is also discussed. This type of analysis, containing a mixture of uncertainty types, requires a theory that is broad enough to have as specific instances all the uncertainty types of interest. To this end, the theory of interval-valued probability measures (IVPMs) (Weichselberger, 2000; Jamison and Lodwick, 2004; Lodwick and Jamison, 2006), and the theory of clouds (Neumaier, 2004a) are presented. The problem of mathematical analysis on distributions is twofold. First, it requires knowing how to do arithmetic over general distributions. Second, it requires knowing how to evaluate a function of distributions in such a way that when given the input (domain) of lower and upper distribution bounds containing the distribution, the result or output (range) will be verified. The relationship between convolutions and distribution arithmetic are mentioned.

Convolutions are the definitions of distribution arithmetic and are usually computationally intractable for complex expressions. Our interest is in relating the theoretical foundations (Springer, 1979; Kaplan, 1981), in particular, copulas (Nelsen, 1995), to bounds on distributions resulting from binary operations and their relationship to interval and fuzzy arithmetic (Dubois and Prade, 1987a; Williamson and Downs, 1990a; Williams, 1990b). Arithmetic operations, of course, are binary operations.

The material presented in this section is not an exhaustive survey, but consists of a few methods that are more tractable. We have already discussed enclosure and verification associated with intervals that can be considered a type of uncertainty. In the context of fuzzy, possibilistic, and probabilistic uncertainty, lower and upper distributions need to be obtained or constructed (Jamison and Lodwick, 2002). For fuzzy uncertainty, a fuzzy set as a crisp in $\mathbb{R}^2$ may be thought as enclosing the uncertainty that a quantity $x$ is (or is in) $A$, with the *x-axis* being the lower distribution and the membership function being the upper distribution so that $0 \leq \textit{uncertainty}_A(x) \leq \mu_A(x)$. This was pointed out by Dubois *et al.* (1997). From this point of view, a fuzzy set membership function loses information at the lower bound. For possibilistic uncertainty, possibility and its dual, necessity, pairs can be developed to provide a "tighter" set of bounds on the uncertainty, if constructed appropriately (Jamison and Lodwick, 2002).

The question of a general theory containing all uncertainties of interest to this monograph (interval, fuzzy, possibility, probability) was dealt with by Jamison and Lodwick (2002). In particular, Jamison and Lodwick (2004), Lodwick and Jamison (2003b, 2006) consider IVPMs as the theory behind the methods for obtaining lower and upper bounds on all of the uncertainties. Clouds (see Neumaier, 2004a) are also enclosures of uncertainty that include the uncertainty associated with probabilities, intervals, and fuzzy sets, as well as possibility. Clouds can also be viewed as IVPMs (Lodwick and Jamison, 2006). Weichselberger (2000) develops interval-valued probabilities that are

used for bounds on probability distributions. His work is extended to explicitly include intervals, fuzzy, possibility, and clouds by Jamison and Lodwick (2006) and Lodwick and Jamison (2006). When one has valid bounds on entities, whose values are characterized by probabilistic, fuzzy set, and/or possibilistic uncertainties, and these values are used in a mathematical model described by a function to verify the result, then propagation of the lower and upper distribution in a way that guarantees that the derived lower and upper bounds enclose the uncertainty, is verification. The exposition begins with arithmetic.

### A. *Distribution Arithmetic*

This section begins with several approaches to distribution arithmetic that attempt to be computationally tractable. Once distribution arithmetic is developed, how to analyze algebraic functions by means of a new extension principle is presented in the section on IVPM. The problem being addressed, in the context of probability theory, is, given a real-valued function

$$f : \mathbb{R}^n \to \mathbb{R},$$

compute the cumulative distribution function of

$$Y = f(X), \tag{58}$$

where $X$ is a vector of independent random variables. The cumulative distribution function of $Y$ is denoted $F_Y$. It is first assumed that, $f$, as a real-valued function, is continuous and monotonically increasing in each variable. It is a theoretically simple adjustment (although it may be computationally complex) to consider functions that increase in some variables and decrease in others. It is also assumed that each random variable $X_i$ has a finite closed and bounded support $[\underline{s}_i, \bar{s}_i]$ with a marginal cumulative distribution function denoted $F_{X_i}$. That is, the support of $X_i$ is

$$\text{supp}(X_i) = [\underline{s}_i, \bar{s}_i] = \left[F_{X_i}^{-1}(0), F_{X_i}^{-1}(1)\right]. \tag{59}$$

Given the assumption that $F_{X_i} : [\underline{s}_i, \bar{s}_i] \to [0, 1]$ is one-to-one, onto and strictly increasing, it is invertible. Let $F_{X_i}^{-1}$ denote this inverse, and $F_Y : [\underline{s}_Y, \bar{s}_Y] \to [0, 1]$ denote the cumulative distribution function of $Y = f(X)$, which is also invertible since $f$ is monotonically increasing. The support of each random variable is a closed interval. Clearly, $Y$ is contained in a closed and bounded interval given that the support of each $X_i$ is a closed and bounded interval and $f$ is continuous.

The methods outlined in this section are useful in:

1. *Simulations* where a resulting closed form is desired [Eq. (58)]. Moreover, circumventing a Monte Carlo approach altogether may be desirable, given the complexity of the problem.
2. *Optimization under uncertainty* (Jamison and Lodwick, 2006; Lodwick and Jamison, 2006).
3. *Risk analysis* (Kaplan, 1981; Berleant and Goodman-Strauss, 1998; Jamison *et al.*, 2002; Ferson *et al.*, 2003; Berleant and Zhang, 2004).
4. *Enclosure/verification* where the idea is to enclose distributions between lower and upper envelopes (Berleant, 1993; Jamison and Lodwick, 2002, 2006; Lodwick and Jamison, 2006; Moore, 1984; Williamson and Downs, 1990a; Williams, 1990b).

There are a variety of approaches to distribution arithmetic. The three types presented here are: (1) the *interval convolution* approach based on Williamson and Downs (1990a), Williams (1990b); (2) the *interval histogram* approaches, which are based on histograms (Moore, 1984; Berleant, 1993; Tonon, 2004); and (3) the *inverse probability* approach based on Lodwick and Jamison (2003b) and Olsen (2005). These methods are developed for cumulative distribution functions. To extend these methods to possibility, a method to construct quantitative possibility and necessity distribution pairs such as Jamison and Lodwick (2002) must be used. Given an appropriately created possibility and necessity pair, the enclosure of the pair is obtained by taking the upper function of the envelope for the possibility and the lower function of the envelope for the necessity. That is, if the constructed possibility function is $p(x)$ and the constructed necessity function is $n(x)$, and these functions are derived using Jamison and Lodwick (2002), then, using methods outlined below, envelopes on $n(x)$ and $p(x)$ are constructed, so that:

$$g_p(x) \leq p(x) \leq h_p(x),$$
$$g_n(x) \leq n(x) \leq h_n(x).$$

The enclosure of possibility and necessity pair becomes,

$$g_n(x) \leq n(x) \leq p(x) \leq h_p(x).$$

Thus, it is straightforward to extend the methods outlined in this section to possibility and necessity pairs, and thus it will not be treated separately.

### 1. *Interval Convolution Methods*

Convolutions used in defining distribution arithmetic were known from the start. Following the development of Kaplan (1981), Williamson and Downs (1990a), Williams (1990b), given a general binary relationship,

$$W = f(X, Y), \tag{60}$$

where $X$ and $Y$ are independent random variables, the probability density function for the random variable resulting from the operation is given by

$$p_W(w) = \int_{-\infty}^{\infty} p_X(x) p_Y\big(f^{-1}(w,x)\big) \frac{\partial}{\partial w}\big(f^{-1}(w,x)\big)\, dx. \tag{61}$$

Here, $f^{-1}(w,x)$ is assumed to exist and is the unique-valued $y$ for which

$$w = f\big(x, f^{-1}(w,x)\big).$$

In particular (Williams, 1990b),

**Addition**

$$p_W(w) = \int_{-\infty}^{\infty} p_X(x) p_Y(w-x)\, dx. \tag{62}$$

**Subtraction**

$$p_W(w) = \int_{-\infty}^{\infty} p_X(x) p_Y(w+x)\, dx. \tag{63}$$

**Multiplication**

$$p_W(w) = \int_{-\infty}^{\infty} \frac{1}{|x|} p_X(x) p_Y\left(\frac{w}{x}\right) dx. \tag{64}$$

**Division**

$$p_W(w) = \int_{-\infty}^{\infty} |x| p_X(x) p_Y(wx)\, dx. \tag{65}$$

If the distributions are dependent, then

$$p_W(w) = \int_{-\infty}^{\infty} p_X(x) p_Y\big(f^{-1}(w,x)|x\big) \frac{\partial}{\partial w}\big(f^{-1}(w,x)\big)\, dx. \tag{66}$$

For example, in the addition of two dependent distributions,

$$p_W(w) = \int_{-\infty}^{\infty} p_X(x) p_Y\big((w-x)|x\big)\, dx. \tag{67}$$

The numerical calculation of the arithmetic defined by its convolution is often computationally intensive, especially in the presence of dependencies whose precise nature (in a closed form) is usually missing. The idea is to compute with just the information associated with the marginal distribution. The error in computing as though the random variables are independent is called *dependency error* by Williamson and Downs (1990a) and Williams (1990b).

**a. Williamson and Downs.** Williamson and Downs (1990a, 1990b) use enclosure to deal with dependency bounds associated with distribution arithmetic and functions of distributions. Their research not only deals with enclosure but also deals directly with dependencies. Note that constraint interval arithmetic (Lodwick, 1999) and gradual numbers (Fortin *et al.*, 2006) also make explicit the issue of dependencies. In particular, Williamson and Downs use Fréchet bounds to enclose the joint cumulative distribution as follows:

$$\max\{F_X(x) + F_Y(y) - 1, 0\} \leq F_{XY}(x, y) \leq \min\{F_X(x), F_Y(y)\} \quad \forall x, y \in \mathbb{R}. \tag{68}$$

Clearly, Eq. (68) is related to what the fuzzy literature calls a *t-norm* (Klir and Yuan, 1995). The bounds of Eq. (68) are tight since they are met in the cases of positive or negative dependence. Williams and Downs go on to use copulas that allow the Fréchet bounds to be written in a perhaps more useful way and whose definition is as follows.

***Definition 16*** (Nelsen, 1995). A (two-dimensional) *copula* is a function

$$C : I \times I \to I = [0, 1],$$

with the following properties:

(1) $C(0, t) = C(t, 0) = 0$ and $C(1, t) = C(t, 1) = t, \forall t \in I$,
(2) $C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_i, v_1) \geq 0, \forall u_1, u_2, v_1, v_2 \in I$,

where $u_1 \leq u_2$ and $v_1 \leq v_2$.

***Remark.*** $C$ is nondecreasing in each variable and continuous (since it satisfies a Lipschitz condition) $|C(u_2, v_2) - C(u_i, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|$.

***Definition 17*** (Williams, 1990b). A *connecting copula* for the random variables $X$ and $Y$ is the copula

$$C_{XY}(u, v) = F_{XY}\big(F_X^{-1}(u), F_Y^{-1}(v)\big). \tag{69}$$

***Remark.*** It is clear that a connecting copula is a copula. Moreover,

$$F_{XY}(u, v) = C_{XY}\big(F_X(u), F_Y(v)\big). \tag{70}$$

Two other copulas are

$$W(a, b) = \max\{a + b - 1, 0\} \tag{71}$$

and

$$M(a, b) = \min\{a, b\}. \tag{72}$$

Williams (1990b) mentions that, "The copula $C_{XY}$ contains all information regarding the dependency of $X$ and $Y$. If $C_{XY}(x, y) = xy$, then $X$ and $Y$ are independent."

Given a binary operation on two random variables $Z = X \circ Y$ ($\circ$ is the binary operation) that is monotonic in both variables (places) where the marginal distributions $F_X$ and $F_Y$ are known, bounds on $F_Z$ are sought (called dependency bounds by Williamson and Downs; in the language of this monograph, dependency bounds are *enclosures*). That is,

$$ldb(F_X, F_Y, \circ)(z) \leq F_Z(z) \leq udb(F_X, F_Y, \circ)(z), \quad \forall z \in \mathbb{R},$$

where *ldb* and *udb* denote the lower dependency bound and the upper dependency bound, respectively. It has been shown by Frank *et al.* (1987) that for certain binary operations that include $\circ \in \{+, -, \times, \div\}$,

$$ldb(F_X, F_Y, \circ)(z) = \sup_{z=x\circ y} W\big(F_X(x), F_Y(y)\big)$$

and

$$ldb(F_X, F_Y, \circ)(z) = \inf_{z=x\circ y} W^d\big(F_X(x), F_Y(y)\big)$$

where $W^d$ is called the *dual copula* given by

$$W^d(x, y) = x + y - W(x, y).$$

Williamson and Downs go on to show how to compute the lower and upper dependency bounds that have order $O(n^2)$ where $n$ is the number of discretization of the distributions $F$. Their approach has been implemented by Ferson (2002), Ferson *et al.* (2003), and Ferson and Kreinovich (2005), and is similar to quantile arithmetic discussed previously (Dempster, 1969, 1974; Nickel, 1969).

2. *Interval Histogram Methods*

This section presents three interval histogram methods. Broadly speaking, interval histogram methods are based on partitioning the supports [Eq. (59)], computing histograms on the partition, and using interval arithmetic to compute an approximation to $F_Y$.

**a. R.E. Moore.** The precursor and basis of all interval histogram methods, the two inverse probability methods (as well as the interval convolution method above), is Moore's (1984) article. The method outlined by Moore *approximates* the cumulative distribution of a random variable and uses this to approximate the cumulative distribution of a function of random variable [Eq. (58)]. There is no attempt to enclose the correct cumulative distribution, although subsequent approaches, in particular, Berleant (1993), Lodwick and Jamison (2003b), and Olsen (2005), do. Given Eq. (58) along with the stated assumption, Moore constructs an approximation to the cumulative distribution $F_Y$ of the function $Y = f(X)$ in the following way.

***Algorithm IV.1.*** Moore's Method (Moore, 1984)

*Step 1. Partition*—The support of each random variable, $X_i$, is subdivided into $N_i$ subintervals, usually of equal width.

*Step 2. Compute the probability of the partition*—The probability (histogram) for each random variable, on each of its subintervals, is computed from the given probability density function.

*Step 3. Compute the approximate value of the function on each subinterval*—Use interval analysis on each of the $K = \prod_{i=1}^{n} N_i$ combinations of the subintervals. Each instantiation yields an interval where the probability assigned to the interval is the product of marginal probabilities, and Moore assumes that the random variables are independent.

*Step 4. Order the resultant intervals and subdivide overlapping segments*—The intervals obtained in Step 3 are ordered with respect to their left end points. Any overlapping intervals are subdivided.

*Step 5. Compute the probabilities on the overlapping segments*—Probabilities are assigned to the subdivided intervals in proportion to their length. That is, Moore (1984) assumes that if an interval is subdivided into say, three parts, each part of the interval receives one third of the probability associated with that interval.

*Step 6. Assemble into one cumulative distribution function*—The probability of an overlapping set of subintervals is the sum of the probabilities on each subinterval. Starting with the left-most interval, the range value of the cumulative distribution at the left end point is 0, and the right end point is probability assigned to that subinterval. Linear interpolation between the

two range values is used as an approximation. The cumulative distribution at the right end point of the second subinterval is the sum of the probability of the first subinterval and second subinterval. The process is continued.

***Example.*** Suppose $Y = X_1 + X_2$, where

$$X_1 = X_2 = \begin{cases} 1 & x \in [0, 1] \\ 0 & \text{otherwise,} \end{cases}$$

that is, $X_1$ and $X_2$ are uniform independent random distributions on $[0, 1]$, denoted $U[0, 1]$. The problem is to compute, using Moore's approach, the cumulative distribution of $Y$.

*Step 1. Partition*—$X_1$ and $X_2$ are partitioned into two subintervals each, $X_1 = X_{11} \cup X_{12} = [0, \frac{1}{2}] \cup [\frac{1}{2}, 1]$, and $X_2 = X_{21} \cup X_{22} = [0, \frac{1}{2}] \cup [\frac{1}{2}, 1]$, resulting in a partition consisting of four boxes: (1) $[0, \frac{1}{2}] \times [0, \frac{1}{2}]$; (2) $[\frac{1}{2}, 1] \times [0, \frac{1}{2}]$; (3) $[0, \frac{1}{2}] \times [\frac{1}{2}, 1]$; (4) $[\frac{1}{2}, 1] \times [\frac{1}{2}, 1]$ (Figure 1).

*Step 2. Compute the probability of the partition*—Since the probability on each subinterval is $\frac{1}{2}$, and the random variables are assumed to be independent, the probability on each box is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.

*Step 3. Compute the approximate value of the function on each subinterval*—Use interval arithmetic (or constraint interval arithmetic) to compute the value of $Y$ on each box.

For box (1), $Y_1 = [0, \frac{1}{2}] + [0, \frac{1}{2}] = [0, 1]$, where the probability for $Y_1$ is $\frac{1}{4}$.

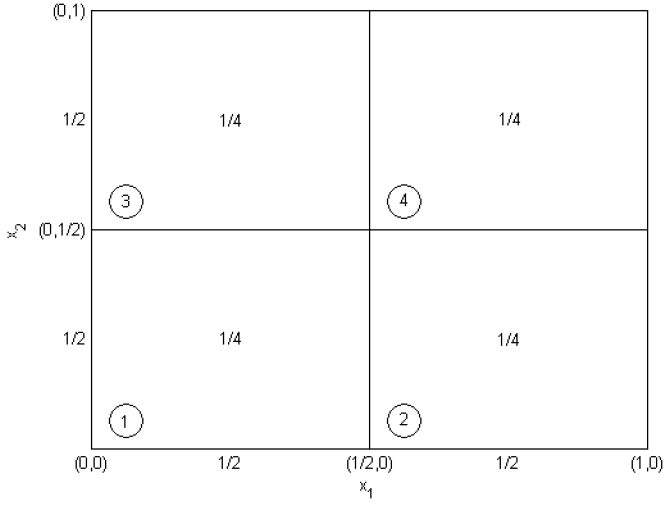


FIGURE 1. Partition of $[0, 1] \times [0, 1]$.

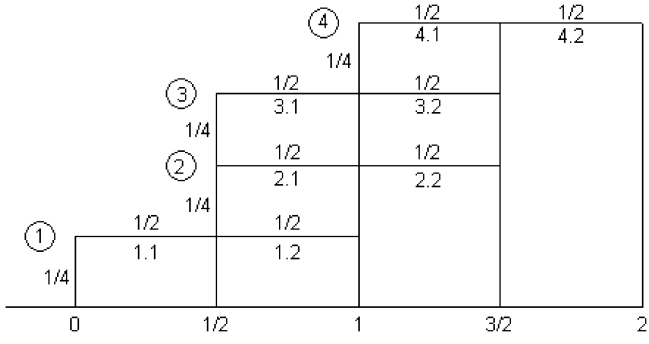


FIGURE 2. Resultant intervals and overlaps.

For box (2), $Y_2 = [\frac{1}{2}, 1] + [0, \frac{1}{2}] = [\frac{1}{2}, \frac{3}{2}]$, where the probability for $Y_2$ is $\frac{1}{4}$.

For box (3), $Y_3 = [0, \frac{1}{2}] + [\frac{1}{2}, 1] = [\frac{1}{2}, \frac{3}{2}]$, where the probability for $Y_3$ is $\frac{1}{4}$.

For box (4), $Y_4 = [\frac{1}{2}, 1] + [\frac{1}{2}, 1] = [1, 2]$, where the probability for $Y_4$ is $\frac{1}{4}$.

*Step 4. Order the resultant intervals and subdivide overlapping segments*—From the computations, the intervals are ordered according to their left end points as follows: $Y_1$, $Y_2$, $Y_3$, $Y_4$, with distinct (overlapping) subintervals of $[0, \frac{1}{2}]$—one subinterval; $[\frac{1}{2}, 1]$—three subintervals; $[1, \frac{3}{2}]$—three subintervals; and $[\frac{3}{2}, 2]$—one subinterval (Figure 2).

*Step 5. Compute the probabilities on the overlapping segments*—Since $[0, \frac{1}{2}]$ came from subdividing $Y_1 = [0, 1]$ in half, the probability on the first subinterval is $\frac{1}{2} \times \frac{1}{4} = \frac{1}{8}$. The assumption that Moore makes is that the proportion of the division of the interval is the probability. In the same manner, there are three overlapping subintervals comprising $[\frac{1}{2}, 1]$ arising from half portions of $Y_1$, $Y_2$, and $Y_3$, each bearing probability of $\frac{1}{2} \times \frac{1}{4} = \frac{1}{8}$, so that the probability on $[\frac{1}{2}, 1]$ is $\frac{3}{8}$. In the same manner, the probability on $[1, \frac{3}{2}]$ is $\frac{3}{8}$ and on $[\frac{3}{2}, 1]$ is $\frac{1}{8}$ (Figure 2).

*Step 6. Assemble into one cumulative distribution function*—

$$Y(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{8}, & x = \frac{1}{2} \\ \frac{1}{8} + \frac{3}{8} = \frac{1}{2}, & x = 1 \\ \frac{4}{8} + \frac{3}{8} = \frac{7}{8}, & x = \frac{3}{2} \\ \frac{1}{8} + \frac{7}{8} = 1, & x = 2 \\ 1, & x > 1 \end{cases}$$

with linear interpolation in between (Figure 3).

***Example.*** Suppose $Y = X_1 + X_2$, where $X_1 = X_2$ are normal distributions $X_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, $X_2(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ and we consider that an interval of plus or minus 3 standard deviations (SDs) from the mean value of 0 (Figure 4).

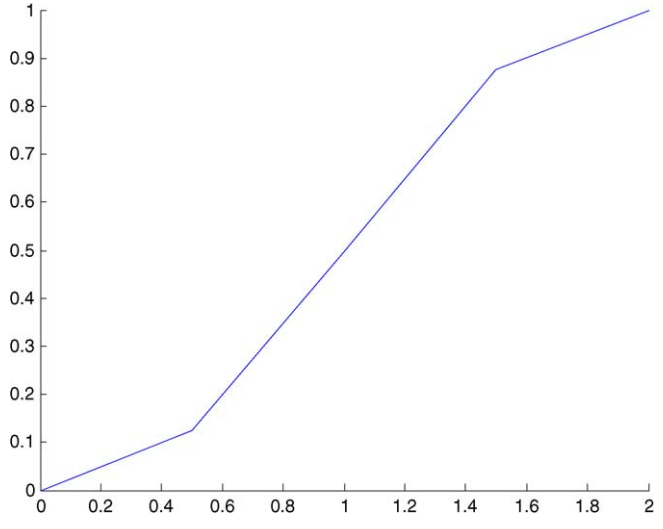


FIGURE 3. Assembled cumulative distribution function.

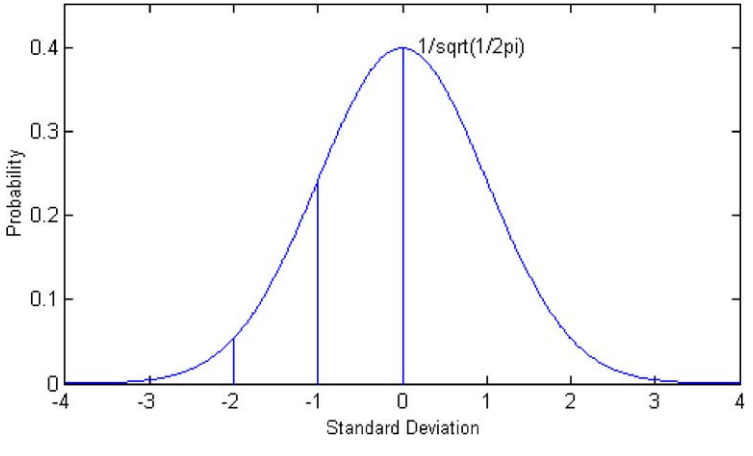


FIGURE 4. Normal distribution divided in units of 1 standard deviation.

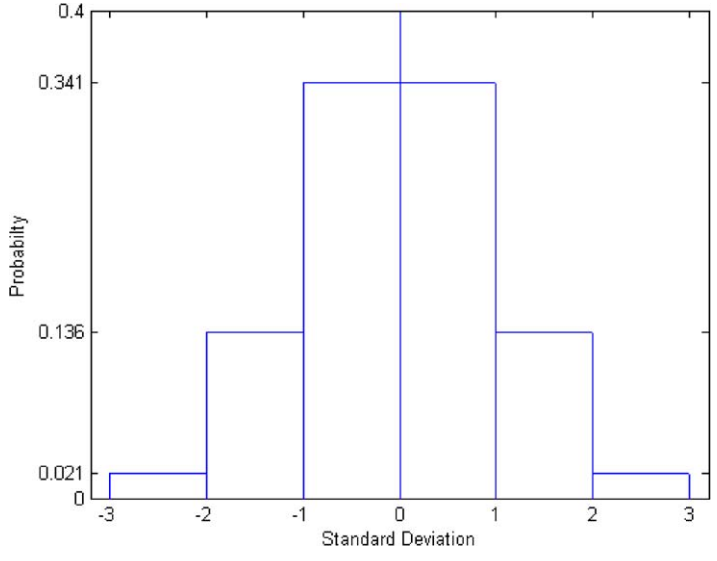


FIGURE 5. Normal distribution in block approximation.

*Step 1. Partition*—$X_1$ and $X_2$ are partitioned into six subintervals each (Figure 5),

$$X_1 = X_{11} \cup X_{12} \cup X_{13} \cup X_{14} \cup X_{15} \cup X_{16}$$
$$X_2 = X_{21} \cup X_{22} \cup X_{23} \cup X_{24} \cup X_{25} \cup X_{26}$$
$$= [-3\sigma, -2\sigma] \cup [-2\sigma, -\sigma] \cup [-\sigma, 0] \cup [0, \sigma] \cup [\sigma, 2\sigma] \cup [2\sigma, 3\sigma],$$

resulting in a partition consisting of 36 boxes (1) $[-3\sigma, -2\sigma] \times [-3\sigma, -2\sigma]$; (2) $[-3\sigma, -2\sigma] \times [-2\sigma, -\sigma]$; (3) $[-3\sigma, -2\sigma] \times [-\sigma, 0]$; (4) $[-3\sigma, -2\sigma] \times [0, \sigma]$; (5) $[-3\sigma, -2\sigma] \times [\sigma, 2\sigma]$; (6) $[-3\sigma, -2\sigma] \times [2\sigma, 3\sigma]$; ..., (36) $[2\sigma, 3\sigma] \times [2\sigma, 3\sigma]$ (see Figure 6).

*Step 2. Compute the probability of the partition*—Since the probability on each subinterval is divided by equal SDs, for $[-3\sigma, -2\sigma]$, $[-2\sigma, -\sigma]$, $[-\sigma, 0]$, $[0, \sigma]$, $[\sigma, 2\sigma]$, and $[2\sigma, 3\sigma]$ where the associated probabilities are 0.021, 0.136, 0.341, 0.341, 0.136, and 0.021, respectively, the probabilities on the 36 boxes are the product of all pairwise combinations of these 6 numbers. That is, for box (1) $0.021 \times 0.021$; for box (2) $0.021 \times 0.136$; for box (3) $0.021 \times 0.341$; for box (4) $0.021 \times 0.341$; for box (5) $0.021 \times 0.136$; for box (6) $0.021 \times 0.021$; ... for box (36) $0.021 \times 0.021$.

*Step 3. Compute the approximate value of the function on each subinterval*—For box (1), $Y_1 = [-3\sigma, -2\sigma] + [-3\sigma, -2\sigma] = [-6\sigma, -4\sigma]$, where the probability for $Y_1$ is 0.000441. For box (2), $Y_2 = [-3\sigma, -2\sigma] +$

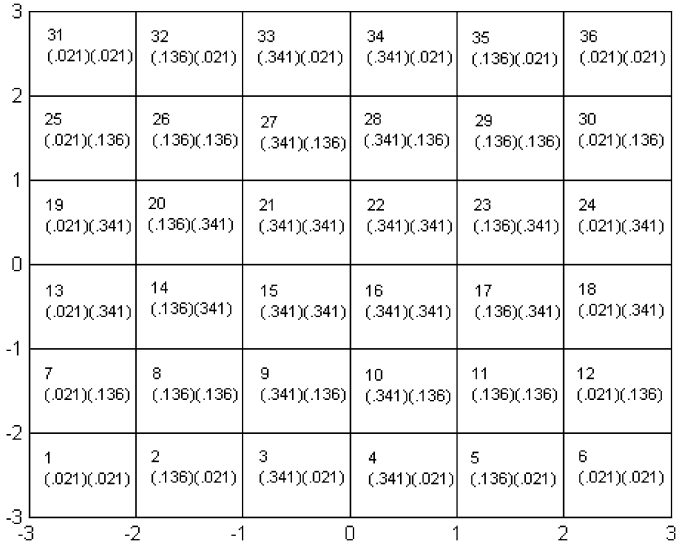


FIGURE 6. Partition of $-3$ standard deviations to 3 standard deviations in the plane.

$[-2\sigma, -\sigma] = [-5\sigma, -3\sigma]$, where the probability for $Y_2$ is 0.002856. For box (3), $Y_3 = [-3\sigma, -2\sigma] + [-\sigma, 0] = [-4\sigma, -2\sigma]$, where the probability for $Y_3$ is 0.007161. For box (4), $Y_4 = [-3\sigma, -2\sigma] + [0, \sigma] = [-3\sigma, -\sigma]$, where the probability for $Y_4$ is 0.007161. For box (5), $Y_5 = [-3\sigma, -2\sigma] + [\sigma, 2\sigma] = [-2\sigma, 0]$, where the probability for $Y_5$ is 0.002856. For box (6), $Y_6 = [-3\sigma, -2\sigma] + [2\sigma, 3\sigma] = [-\sigma, \sigma]$, where the probability for $Y_6$ is 0.000441. For box (7), $Y_7 = [-3\sigma, -2\sigma] + [-2\sigma, -\sigma] = [-5\sigma, -3\sigma]$, where the probability for $Y_7$ is $0.021 \times 0.126 = 0.002646$. Continuing, for box (36), $Y_{36} = [2\sigma, 3\sigma] + [2\sigma, 3\sigma] = [4\sigma, 6\sigma]$, where the probability for $Y_{36}$ is 0.000441.

*Step 4. Order the resultant intervals and subdivide overlapping segments*—From the computations, the intervals are ordered according to their left end points as follows: $Y_1, Y_2, Y_7, Y_3, Y_8, Y_{13}, \ldots, Y_{36}$ with distinct (overlapping) subintervals of $[-6\sigma, 6\sigma]$. $Y_1$ through $Y_{36}$ are subdivided into two subintervals each with a width of 1 SD $\sigma$. The associated probabilities in each box are given in Figure 6. (See also Table 1.)

*Step 5. Compute the probabilities on the overlapping segments*—Since each $Y_i, i = 1, \ldots, 36$, is halved, the probability on each subinterval of $Y_i$ is $\frac{1}{2}$ of the probability on $Y_i$. Thus, we have the following resulting probabilities: For $[-6\sigma, -5\sigma]$, $\frac{1}{2} \times 0.000441 = 0.0002205$. For $[-5\sigma, -4\sigma]$, $\frac{1}{2} \times 0.000441 + \frac{1}{2} \times 0.002856 + \frac{1}{2} \times 0.002646 = 0.0029715$. The remaining intervals are computed in the same manner.

TABLE 1

| | | Units $\sigma$ | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $Y_i$ | $[-6,-5]$ | $[-5,-4]$ | $[-4,-3]$ | $[-3,-2]$ | $[-2,-1]$ | $[-1,0]$ | $[0,1]$ | $[1,2]$ | $[2,3]$ | $[3,4]$ | $[4,5]$ | $[5,6]$ |
| Box 1 | $[-6,-4]$ | 1 | 1 | | | | | | | | | | |
| Box 2 | $[-5,-3]$ | | 1 | 1 | | | | | | | | | |
| Box 3 | $[-4,-2]$ | | | 1 | 1 | | | | | | | | |
| Box 4 | $[-3,-1]$ | | | | 1 | 1 | | | | | | | |
| Box 5 | $[-2,0]$ | | | | | 1 | 1 | | | | | | |
| Box 6 | $[-1,1]$ | | | | | | 1 | 1 | | | | | |
| Box 7 | $[-5,-3]$ | | 1 | 1 | | | | | | | | | |
| Box 8 | $[-4,-2]$ | | | 1 | 1 | | | | | | | | |
| Box 9 | $[-3,-1]$ | | | | 1 | 1 | | | | | | | |
| Box 10 | $[-2,0]$ | | | | | 1 | 1 | | | | | | |
| Box 11 | $[-1,1]$ | | | | | | 1 | 1 | | | | | |
| Box 12 | $[0,2]$ | | | | | | | 1 | 1 | | | | |
| Box 13 | $[-4,-2]$ | | | 1 | 1 | | | | | | | | |
| Box 14 | $[-3,-1]$ | | | | 1 | 1 | | | | | | | |
| Box 15 | $[-2,0]$ | | | | | 1 | 1 | | | | | | |
| Box 16 | $[-1,1]$ | | | | | | 1 | 1 | | | | | |
| Box 17 | $[0,2]$ | | | | | | | 1 | 1 | | | | |
| Box 18 | $[1,3]$ | | | | | | | | 1 | 1 | | | |
| Box 19 | $[-3,-1]$ | | | | 1 | 1 | | | | | | | |
| Box 20 | $[-2,0]$ | | | | | 1 | 1 | | | | | | |
| Box 21 | $[-1,1]$ | | | | | | 1 | 1 | | | | | |
| Box 22 | $[0,2]$ | | | | | | | 1 | 1 | | | | |
| Box 23 | $[1,3]$ | | | | | | | | 1 | 1 | | | |
| Box 24 | $[2,4]$ | | | | | | | | | 1 | 1 | | |
| Box 25 | $[-2,0]$ | | | | | 1 | 1 | | | | | | |
| Box 26 | $[-1,1]$ | | | | | | 1 | 1 | | | | | |
| Box 27 | $[0,2]$ | | | | | | | 1 | 1 | | | | |
| Box 28 | $[1,3]$ | | | | | | | | 1 | 1 | | | |
| Box 29 | $[2,4]$ | | | | | | | | | 1 | 1 | | |
| Box 30 | $[3,5]$ | | | | | | | | | | 1 | 1 | |
| Box 31 | $[-1,1]$ | | | | | | 1 | 1 | | | | | |
| Box 32 | $[0,2]$ | | | | | | | 1 | 1 | | | | |
| Box 33 | $[1,3]$ | | | | | | | | 1 | 1 | | | |
| Box 34 | $[2,4]$ | | | | | | | | | 1 | 1 | | |
| Box 35 | $[3,5]$ | | | | | | | | | | 1 | 1 | |
| Box 36 | $[4,6]$ | | | | | | | | | | | 1 | 1 |

*Step 6. Assemble into one cumulative distribution function—*

$$F_y(x) = \begin{cases} 0 & x \in (-\infty, -6\sigma] \\ 0.0002205 & x = -5\sigma \\ 0.0031921 & x = -4\sigma \\ \vdots & \\ 1 & x \in [6\sigma, \infty) \end{cases}$$

with linear interpolation in between.

**b. Berleant.** Berleant and colleagues' idea, articulated in Berleant (1993), Berleant and Goodman-Strauss (1998), and Berleant and Zhang (2004), was to use the Moore approach (Moore, 1984) (described in the previous subsection), adding enclosure and methods to handle known and unknown dependencies. Williamson and Downs (see the above and Williamson and Downs (1990a); Williams (1990b)) also have this objective. Moreover, the method developed can compute enclosures. They are able to combine interval uncertainty with probabilistic uncertainty, which Monte Carlo cannot do. However, their method does not deal with fuzzy, possibilistic, or cloud uncertainties. The histogram method is first developed for independent variables as follows (we begin with two variables $X$ and $Y$ whose support is assumed to be finite). Assume that $Z = X \circ Y$, where $\circ \in \{+, -, \times, \div\}$. We omit a full discussion of Berleant's approach to enclosure since a complete discussion of the method of Lodwick and Jamison (2003b), in conjunction with Lodwick and Jamison (2006), which provides a complete theoretical framework for mixed uncertainties, is presented in the subsections that follow.

***Algorithm IV.2*** (Berleant, 1993).

*Step 1.* $X$ and $Y$ are discretized using histograms [as was done by Moore (1984)]. Each histogram bar is characterized by both an interval describing its location on the real number line and by a probability mass.

*Step 2.* Obtain the Cartesian product $(X_i, Y_j)$ of all histogram bars describing $X$ and $Y$.

*Step 3.* For all Cartesian products $(X_i, Y_j)$ compute

*Step 3.1.* $Z_{ij} = X_i \circ Y_j$ using interval arithmetic.

*Step 3.2.* Associate with each $Z_{ij}$ the probability $p(Z_{ij}) = p(X_i)p(Y_j)$ according to

$$\begin{aligned} p(x \in X) &\in [0, 1], \\ p(y \in Y) &\in [0, 1], \\ p(z \in Z) &= p(x \in X)p(y \in Y). \end{aligned} \tag{73}$$

*Step 4.* This collection of intervals with associated probabilities is used to construct the probability density function for $Z$ as follows:

*Step 4.1.* Partition $Z$ into intervals on which the histogram for $Z$ will be placed—$Z = Z_1 \cup \cdots \cup Z_K$.

*Step 4.2.* Calculate the mass density that will be placed onto each $Z_k$ as follows:

*Step 4.2.1.* Any $Z_{ij}$ that falls entirely inside the partition gets its associated probability mass assigned to $Z_k$.

*Step 4.2.2.* Any $Z_{ij}$ that overlaps $Z_k$ gets the ratio of the overlapped portion divided by its total assigned to $Z_k$.

*Step 4.2.3.* Drawing the histogram associated with $Z_k$ involves computing the height $h_k$ where

$$h_k = \frac{p(Z_k)}{width(Z_k)}.$$

The process outlined by Berleant and colleagues is precisely Moore's method (given above), except that the histogram for the probability density function is obtained rather than the cumulative distribution function as obtained by Moore. As with Moore, Berleant assumes that ratios of overlaps translate into the same ratio assigned to the probability given the overlap. However, Berleant and colleagues continue their analysis to include known and unknown dependencies. A problem with known dependency is illustrated as follows.

***Example*** (Berleant, 1993). Let

$$X(x) = \begin{cases} \frac{1}{2} & x \in [1, 2] \\ \frac{1}{4} & x \in [2, 4] \\ 0 & \text{otherwise} \end{cases}$$

and

$$Y(x) = \begin{cases} \frac{1}{4} & x \in [2, 3] \\ \frac{1}{2} & x \in [3, 4] \\ \frac{1}{4} & x \in [4, 5] \\ 0 & \text{otherwise.} \end{cases}$$

Assume that $X$ and $Y$ are positively correlated and $Z = XY$.

$$\begin{aligned} x \in [1, 2], &\quad p_{X_1} = \frac{1}{2}, \\ xy \in [2, 6], &\quad p_{X_1Y_1} = 1\frac{1}{4} = \frac{1}{4}, \\ xy \in [3, 8], &\quad p_{X_1Y_2} = \frac{1}{2}\frac{1}{2} = \frac{1}{4}, \\ xy \in [4, 10], &\quad p_{X_1Y_3} = 0\frac{1}{4} = 0, \\ x \in [2, 4], &\quad p_{X_2} = \frac{1}{2}, \end{aligned}$$

$$
\begin{aligned}
xy \in [4, 12], \quad & p_{X_2Y_1} = 0\frac{1}{4} = 0, \\
xy \in [6, 16], \quad & p_{X_2Y_2} = \frac{1}{2}\frac{1}{2} = \frac{1}{4}, \\
xy \in [8, 20], \quad & p_{X_2Y_3} = 1\frac{1}{4} = \frac{1}{4}, \\
y = [2, 3], \quad & p_{Y_1} = \frac{1}{4}, \\
y = [3, 4], \quad & p_{Y_2} = \frac{1}{2}, \\
y = [4, 5], \quad & p_{Y_3} = \frac{1}{4}.
\end{aligned}
$$

The two sets of probabilities $p_{X_1Y_1}$, $p_{X_2Y_3}$ and $p_{X_1Y_3}$, $p_{X_2Y_1}$ are as given above because we are assuming positive correlation. If the two random variables were negatively correlated, then $p_{X_1Y_1}$, $p_{X_2Y_3}$ and $p_{X_1Y_3}$, $p_{X_2Y_1}$ switch as follows:

$$
\begin{aligned}
x \in [1, 2], \quad & p_{X_1} = \frac{1}{2}, \\
xy \in [2, 6], \quad & p_{X_1Y_1} = 0\frac{1}{4} = 0, \\
xy \in [3, 8], \quad & p_{X_1Y_2} = \frac{1}{2}\frac{1}{2} = \frac{1}{4}, \\
xy \in [4, 10], \quad & p_{X_1Y_3} = 1\frac{1}{4} = \frac{1}{4}, \\
x \in [2, 4], \quad & p_{X_2} = \frac{1}{2}, \\
xy \in [4, 12], \quad & p_{X_2Y_1} = 1\frac{1}{4} = \frac{1}{4}, \\
xy \in [6, 16], \quad & p_{X_2Y_2} = \frac{1}{2}\frac{1}{2} = \frac{1}{4}, \\
y \in [8, 20], \quad & p_{X_2Y_3} = 0\frac{1}{4} = 0, \\
y = [2, 3], \quad & p_{Y_1} = \frac{1}{4}, \\
y = [3, 4], \quad & p_{Y_2} = \frac{1}{2}, \\
y = [4, 5], \quad & p_{Y_3} = \frac{1}{4}.
\end{aligned}
$$

To obtain the probability density, compute the ratios of the overlaps.

Arithmetic for random variables with unknown or unspecified dependencies are handled in the following manner. All the intervals resulting from operations remain the same as when the variables are independent. Different dependency relationships imply different assignments of probability masses to the joint distribution cells, yielding different bounding curves. Berleant and colleagues wish to bound the space of all such curves (yielding an enclosure). Berleant calls these bounds that form the enclosure, *dependency bounds*, using the terminology of Williamson and Downs (1990a), Williams (1990b). To derive the dependency bounds, the lowest and highest possible probability for all $z$ belonging to the domain of the resultant random variable $Z$ must be known or computed. To do this, a pair (when two random variables are involved in the computation of $Z$) of joint distribution matrices (hypermatrices when more random variables are involved) that will provide the lowest and highest probabilities is formed. After obtaining these lowest and highest values (outlined next), the overlaps are computed (on the lowest and highest values) as presented above.

***Definition 18*** (Berleant and Goodman-Strauss, 1998). A *joint distribution matrix* for marginal discretizations of two random variables $X$ and $Y$ with bar probability masses $X_i, i = 1, \ldots, m$, and $Y_j, j = 1, \ldots, n$, is the $m \times n$ matrix $0 \leq p_{ij} \leq 1$, such that $x_i = \sum_{j=1}^{n} p_{ij}$, $i = 1, \ldots, m$, and $y_j = \sum_{i=1}^{m} p_{ij}$, $j = 1, \ldots, n$, where $\sum_i \sum_j p_{ij} = 1$, $\sum_i x_1 = 1$, $\sum_j y_j = 1$.

***Example.*** Joint distribution matrix

| | | | | | $x$ |
|---|---|---|---|---|---|
| | 0.04 | 0.02 | 0.03 | 0.01 | 0.10 |
| | 0.02 | 0.08 | 0.04 | 0.01 | 0.15 |
| | 0.02 | 0.05 | 0.08 | 0.10 | 0.25 |
| | 0 | 0.01 | 0.09 | 0.20 | 0.30 |
| | 0.02 | 0.04 | 0.06 | 0.08 | 0.20 |
| $y$ | 0.10 | 0.20 | 0.30 | 0.10 | 1 |

***Definition 19*** (Berleant and Goodman-Strauss, 1998). A *vertex matrix* is a joint distribution matrix such that at least $(m - 1) \times (n - 1)$ of its entries are equal to 0.

***Example.*** Vertex matrix

| | | | | | $x$ |
|---|---|---|---|---|---|
| | 0.10 | 0 | 0 | 0 | 0.10 |
| | 0 | 0.15 | 0 | 0 | 0.15 |
| | 0 | 0 | 0.15 | 0.10 | 0.25 |
| | 0 | 0 | 0 | 0.30 | 0.30 |
| | 0 | 0.05 | 0.15 | 0 | 0.20 |
| $y$ | 0.10 | 0.20 | 0.30 | 0.10 | 1 |

The creation of vertex matrices is a key step in computing dependency bounds (enclosures), as will be seen, and an algorithm that does this is presented next.

***Algorithm IV.3*** [The Cross-Off Algorithm for Creating a Vertex Matrix (Berleant and Goodman-Strauss, 1998)].

Input: $x = x_1, \ldots, x_m$, $y = y_1, \ldots, y_n$

*Step 1.* Assign $p_{ij} = *$ to the joint distribution matrix.
*Step 2.* Set $x'_i = x_i$, $i = 1, \ldots, m$, and $y'_j = y_j$, $j = 1, \ldots, n$.
*Step 3.* Repeat until all * are replaced by numbers.

*Step 3.1.* Let $x'_I = \max_i x'_i$, $y'_J = \max_j y'_j$.
*Step 3.2.* If $x'_I \geq y'_J$, $p_{IJ} = y'_J$, and for each $p_{iJ} = *$, set $p_{iJ} = 0$.
*Step 3.3.* If $x'_I \leq y'_J$, $p_{IJ} = x'_I$, and for each $p_{Ij} = *$, set $p_{Ij} = 0$.
*Step 3.4.* $x'_I = x'_I - \min\{x'_I, y'_J\}$.
*Step 3.5.* $y'_J = y'_J - \min\{x'_I, y'_J\}$.

The authors further show that, given two discrete (discretized) random variables, $X$, and $Y$, the cross-off algorithm produces a vertex matrix after no more than $n + m - 1$ iterations. Let $A$ denote a matrix that is a subset of the entries of the joint distribution matrix. That is, let

$$A \subseteq \{p_{ij}, i = 1, \ldots, m, j = 1, \ldots, n\}.$$

In what follows, the matrix $A$ will be chosen to include each entry whose associated interval's lower bound is at or below $z$, at which the probability of the result of interval arithmetic operation is to be maximized to get the upper dependency bound at $z$ or whose associated interval's high bound is at or above the number $z$, at which the probability is to be minimized to get the lower dependency bound at $z$. This is most easily understood by an example. Suppose the joint distribution matrix is given for the negative correlation example and $z = xy = 4$. Then, $A$ consists of $p_{11} = 0$, $p_{12} = \frac{1}{4}$, $p_{13} = \frac{1}{4}$,

and $p_{21} = \frac{1}{4}$ since $4 \in [2, 6]$, $4 \in [3, 8]$, $4 \in [4, 10]$, and $4 \in [4, 12]$, respectively, and the $p_{ij}$ of $A$ are those that correspond to these intervals in the matrix. If $z = xy = 8$ is easily checked that matrix $A$ consists of $p_{12}$, $p_{13}$, $p_{21}$, $p_{22}$, and $p_{23}$. To complete the rest of the matrix (given this $A$), the entries are maximized subject to the condition that the sums of the rows and columns add up to the values of the given marginals. By maximizing the sums $P$ of the missing entries of $A$, we minimize the sum of $1 - P$ for the complement of matrix $A$. How the missing entries of matrix $A$ are maximized is discussed next.

***Definition 20*** (Berleant and Goodman-Strauss, 1998). Given two joint distribution matrices $D$ and $D'$ with entries $\{p_{ij}\}$ and $\{p'_{ij}\}$, respectively, if

$$\sum_{(i,j)\in A} p_{ij} > \sum_{(i,j)\in A} p'_{ij},$$

then it is said that $D$ is *higher* than $D'$ ($D'$ is *lower* than $D$).

Given this definition, it is clear that to compute dependency bounds, the lowest and highest matrices ($\underline{D}$ and $\overline{D}$), consistent with the dependencies as defined by the marginals of $X$ and $Y$, are sought.

***Definition 21*** (Berleant and Goodman-Strauss, 1998). Two vertex matrices are said to be *adjacent* if they have at least $(m-1)(n-1)-1$ zeros in common among their entries.

The algorithm to be developed computes $\overline{D}$ (respectively $\underline{D}$) by repeatedly moving from one adjacent vertex matrix to another. Berleant and Goodman-Strauss (1998) prove (see their Lemma 6.2, page 158) that given any vertex matrix $D_0$, there exists a chain of vertex matrices $D_0, D_1, D_2, \ldots, D_k, \ldots, D_K = \overline{D}$ such that each vertex matrix in the chain is adjacent to and higher than the preceding vertex matrix. Moreover, they also prove (see their Lemma 2.3, page 158) that once $(m-1)(n-1)$ entries of the joint distribution matrix such that all the entries of an entire row and an entire column are not specified, then all of the missing entries are determined. All of this leads to an algorithm that computes $\overline{D}$ as follows:

***Algorithm IV.4.*** [Dependency Bounds Berleant and Goodman-Strauss (1998, pages 158, 159)]

*Step 1.* Given a distribution matrix, use the cross-off algorithm to generate an initial vertex matrix $D_0$ and set $k := 0$.

*Step 2.* Loop.

*Step 2.1.* Specify the locations of the zeros for each potential adjacent vertex matrix.

*Step 2.2.* For each potential adjacent vertex matrix, apply the proof to Lemma 2.3 (Berleant and Goodman-Strauss, 1998) to find the values of the remaining entries. Discard any matrix not satisfying $0 \leq p_{ij} < 1$.

*Step 2.3.* Compute the height of all the remaining adjacent vertex matrices as high or higher than any other adjacent matrix.

*Step 2.4.* If $D_{k+1}$ is higher than $D_k$, then set $k := k+1$ and repeat the loop. Or else, set $\overline{D} = D_k$ and stop.

The computation of $\underline{D}$ is done similarly.

***Example*** (Positive correction example). $z = xy = 4$, where

$$
\begin{aligned}
x \in [1,2], &\quad p_{X_1} = \frac{1}{2},\\
xy \in [2,6], &\quad p_{11}\frac{1}{4},\\
xy \in [3,8], &\quad p_{12} = \frac{1}{4},\\
xy \in [4,10], &\quad p_{13} = 0,\\
x \in [2,4], &\quad p_{X_2} = \frac{1}{2},\\
xy \in [4,12], &\quad p_{21} = 0,\\
xy \in [6,16], &\quad p_{22} = \frac{1}{4},\\
xy \in [8,20], &\quad p_{23} = \frac{1}{4},\\
y = [2,3], &\quad p_{Y_1} = \frac{1}{4},\\
y = [3,4], &\quad p_{Y_2} = \frac{1}{2},\\
y = [4,5], &\quad p_{Y_3} = \frac{1}{4},
\end{aligned}
$$

and

$$
A = \begin{bmatrix} 0 & \frac{1}{4} \\ \frac{1}{2} & * \\ 0 & * \end{bmatrix}, \quad A' = \begin{bmatrix} \frac{1}{4} & 0 \\ \frac{1}{4} & * \\ 0 & * \end{bmatrix}.
$$

$$
\sum p_{ij} = \frac{3}{4} > \sum p'_{ij} = \frac{1}{2}.
$$

Building lower and upper dependency bounds requires the analysis of all $z \in Z = [2, 20]$ for this example. To find the upper dependency bound, the following problem is solved:

$$\max \sum_{(i,j)\in A} p_{ij}$$

subject to

$$\sum_{j=1}^{n} p_{ij} = x_i, \quad i = 1, \ldots, m, \qquad \sum_{i=1}^{m} p_{ij} = y_j, \quad j = 1, \ldots, n.$$

Connecting the maxima for each $z$, we form the upper dependency bound. Similarly, the same procedure is used to compute the lower dependency bound except a linear program that minimizes is used. There exists an uncountably infinite number of $z \in Z = [2, 20]$. However, for intervals, there are a finite number of transitions at which we need to compute the $z$, and these occur at the intersection of our subintervals. For this example, the transitions are at $z = 2, 3, 4, 6, 8, 10, 12, 16, 20$.

**c. Tonon.** Tonon (2004), like Berleant and Goodman-Strauss (1998), uses Moore's approach (Moore, 1984) to bracket the cumulative distribution of a function of random variables using the theory of random sets (Dubois and Prade, 1991) instead of linear programming for enclosure. Given $D \subseteq \mathbb{R}^n$, $P(D)$, the power set of $D$, a *basic probability measure function*, is a set-valued function $m : P(D) \to [0, 1]$ where

$$\forall A \in P(D), \quad m(A) \geq 0,$$
$$m(\emptyset) = 0,$$
$$\sum_{A\in P(D)} m(A) = 1.$$

***Definition 22*** (Demster, 1967; Dubois and Prade, 1991; Klir and Yuan, 1995; Shafer, 1987). If $m(A_\alpha) > 0$, $A_\alpha \in D$, $\alpha \in \Im$, an indexing set, $A_\alpha$ is called a *focal element*. A *random set* is the pair $(\mathcal{F}, p)$ where $\mathcal{F}$ is the set of all focal elements of $D$. For a specific pair $(\mathcal{F}, p)$, $A_\alpha \in \mathcal{F} \Rightarrow m(A_\alpha) > 0$, $\sum_{A_\alpha\in\mathcal{F}} m(A_\alpha) = 1$. For $E \in P(D)$,

$$\begin{aligned} Bel(E) &= \sum_{A_\alpha\subseteq E, A_\alpha\in\mathcal{F}} m(A_\alpha) \leq \Pr(E) \\ &\leq Pl(E) = \sum_{A_\alpha\cap E\neq\emptyset, A_\alpha\in\mathcal{F}} m(A_\alpha). \end{aligned} \tag{74}$$

*Bel* is called the *belief measure* and *Pl* is called the *plausibility measure*.

For random sets, a measure is constructed on subsets of $D$ (not singletons), and these are the outcomes of subsets $A_i \subseteq D$, which are the observations indicating an event somewhere in $A_i$. There is no reference made to a probability of a particular point in $A_i$. Any probability, for which $p(A_i) = m(A_i)$, would be consistent with the basic probability measure function $m$. If the focal elements are all singletons, then, of course, the basic probability measure defines a probability distribution and

$$\begin{aligned} Bel(E) &= \sum_{A_\alpha \subseteq E, A_\alpha \in \mathcal{F}} m(A_\alpha) \\ &= \Pr(E) = Pl(E) = \sum_{A_\alpha \cap E \neq \emptyset, A_\alpha \in \mathcal{F}} m(A_\alpha). \end{aligned}$$

Since $\sum_{A_\alpha \in \mathcal{F}} m(A_\alpha) = 1$, we have

$$\begin{aligned} Bel(E) &= 1 - Pl\big(E^C\big) \\ Pl(E) &= 1 - Bel\big(E^C\big). \end{aligned}$$

For $D \subseteq \mathbb{R}$, each connected focal element, $A_\alpha$, is an interval. For this case, two cumulative distributions can be computed:

$$\underline{F}(u) = Bel\{u' \in \mathbb{R} \mid u' \leq u\} = \sum_{A_\alpha \in \mathcal{F}, u \geq \sup(A_\alpha)} m(A_\alpha),$$

$$\overline{F}(u) = Pl\{u' \in \mathbb{R} \mid u' \leq u\} = \sum_{A_\alpha \in \mathcal{F}, u \geq \inf(A_\alpha)} m(A_\alpha).$$

Note that for any $A_\alpha = [\underline{a}_\alpha, \bar{a}_\alpha]$ such that $\underline{a}_\alpha \leq u$ implies that $A_\alpha \cap (-\infty, u] \neq \emptyset$, and $\bar{a}_\alpha \leq u$ implies that $A_\alpha \subseteq (-\infty, u]$. An expectation is defined as follows.

***Definition 23.*** For all connected focal elements $A_\alpha \in \mathbb{R}$, the *expectation* $\mu$ is

$$\mu = \left[\sum_\alpha m(A_\alpha) \inf(A_\alpha), \sum_\alpha m(A_\alpha) \sup(A_\alpha)\right].$$

The extension principle for random sets is defined next.

***Definition 24.*** Let $y = f(u)$, $f : u \in D \to Y \subseteq \mathbb{R}$. The image of $(\mathcal{F}, p)$ is $(\Re, \rho)$ where $\Re = \{R = f(A_\alpha), A_\alpha \in \mathcal{F}\}$ and $\rho(R) = \sum_{A_\alpha:\ R = f(A_\alpha)} m(A_\alpha)$. Of course, $\rho(R) = 0$ if there exists an $A_\alpha$ such that $R = f(A_\alpha)$. Clearly, $\rho(R) \geq 0$.

Given the above definitions and observations, lower, $\underline{F}$, and upper, $\overline{F}$, bounds on the cumulative distribution, $F$, $\underline{F} \leq F \leq \overline{F}$ are computed as follows. Let $y = f(\vec{\mathbf{x}})$, $m$ a basic probability measure function, and $\{A_i\}$ a partition of nonzero diameter (hence a focal element) such that $D = \bigcup_{i=1}^{N} A_i$.

$$\underline{F}(y) = \sum_{A_i :, y \geq \sup f(A_i)} m(A_i), \tag{75}$$

$$\overline{F}(y) = \sum_{A_i :, y \geq \inf f(A_i)} m(A_i). \tag{76}$$

It is clear by construction that $\underline{F}(y) \leq F(y) \leq \overline{F}(y)$. What the above lower and upper cumulative distribution functions say is that, if $y = f(\vec{\mathbf{x}})$, the range value, is greater than or equal to the maximum over the $i$th partition, $A_i$, then the probability measure of that partition is added to the lower cumulative distribution function. In a similar way, the upper cumulative distribution is formed. The actual value of the distribution inside the partition may be unknown. Tonon proceeds to use the "vertex method," which is a simple way to evaluate the function only at the vertices of the partition (box) $A_i$, where he assumes that the function $f$ is monotonic in each variable, thereby obtaining the minimum and maximum values of the function over this partition, which occur at the vertices (given the monotonicity assumption). This cuts down on the complexity of computing the global optimum. For nonmonotonic functions, one may apply interval global optimization techniques whose overestimation is minimal given partitions (boxes) that have small diameters.

3. *Inverse Probability Method*

Inverse probability is a phrase coined by Olsen (2005) to denote that partitions are computed from the inverse distribution. Jamison and Lodwick (2003b, 2004) partition using the inverse of the marginal cumulative distributions for equally spaced grid points so that the boxes are of equal probability, as does Olsen. Olsen (2005), however, focuses on the development of an efficient algorithm for distribution arithmetic. In particular, the overestimation generated by the enclosure developed is smaller so that the method potentially converges in fewer steps.

**a. Jamison and Lodwick.** The problem addressed by Jamison and Lodwick (2003b, 2004, 2006) is Eq. (58), one for which the function of distributions possesses the associated articulated assumptions given above.

*The Approach.* Given the vector of continuous random variables, $\vec{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$, the joint cumulative distribution function at $\vec{x} = (x_1,$

$\ldots, x_n)^{\mathrm{T}}$, $F_X(\vec{x})$, from the left is

$$F_X(\vec{x}) = prob(X_1 \leq x_1, \ldots, X_n \leq x_n),$$

and, from the right, the joint cumulative distribution function, $G_X(\vec{x})$, is

$$G_X(\vec{x}) = prob(x_1 \leq X_1, \ldots, x_n \leq X_n).$$

Bounds (lower and upper) are constructed, and an approximation between the lower and upper bounds for the cumulative distribution function of $Y$ is computed in three steps. The first step *partitions* the domain $D = \{[\underline{s}_1, \bar{s}_1] \times \cdots \times [\underline{s}_n, \bar{s}_n]\}$ into small boxes. The second step *constructs lower and upper bounds* and an intermediate estimate for the conditional cumulative distribution function of $Y$ for each box of the partition given that $X$ is in that box. The final step *combines* the cumulative distribution functions on each box into the cumulative distribution function over the entire domain.

*Step 1. Partition*—The first step is to construct a partition on the domain, $D$,

$$\begin{aligned} D &= [\underline{s}_1, \bar{s}_1] \times \cdots \times [\underline{s}_n, \bar{s}_n] \\ &= \left[F_{X_1}^{-1}(0), F_{X_1}^{-1}(1)\right] \times \cdots \times \left[F_{X_n}^{-1}(0), F_{X_n}^{-1}(1)\right] \end{aligned}$$

into subboxes of equal probability since this simplifies the derived method. For example, to divide $[F_{X_i}^{-1}(0), F_{X_i}^{-1}(1)]$ into three pieces of equal probability, we use $[F_{X_i}^{-1}(0), F_{X_i}^{-1}(\frac{1}{3})]$, $[F_{X_i}^{-1}(\frac{1}{3}), F_{X_i}^{-1}(\frac{2}{3})]$ and $[F_{X_i}^{-1}(\frac{2}{3}), F_{X_i}^{-1}(1)]$. Overlap is not an issue since the distribution of $X$ is assumed to be continuous. The primary consideration in this process is how partitioning affects the size of the problem. If there are $n$ random variables and each variable $X_i$ is divided into $k_i$ subintervals, then there are $\prod_{i=1}^{n} k_i$ conditional cumulative distribution functions to compute. For this study, we do not explore partitioning strategies. Regardless, it is desirable to minimize the number of subdivisions and subdivide only the variables that influence the results the most (a large partial derivative in the box may be used as an indicator).

***Remark.*** Approximating the actual cumulative distribution function using our approach is most difficult in $[F_Y^{-1}(0), F_Y^{-1}(0 + \delta)]$ and in $[F_Y^{-1}(1 - \delta), F_Y^{-1}(1)]$ since there are no overlaps. The best results for the intermediate approximation are obtained from overlaps where there is, potentially, a cancellation of overapproximations and underapproximations as will be seen.

***Remark.*** Equally spaced partitions in probability require the inverse of $F_{X_i}^{-1}$ to be known in closed form or at least computable. When this is the case, the partitioning strategy proceeds in a straightforward manner. When this is not the case, an additional complexity arises in that the inverse must be approximated.

*Step 2. Construction of Lower and Upper Bounds and an Approximation on Each Partition*—The second step is to construct the bounds and the estimated conditional cumulative distribution function for each box of the partition. Let $[b_1, c_1] \times \cdots \times [b_n, c_n]$ be one such box and let $A$ denote the event $X$ falls in this box. Consider the family of $n$-dimensional boxes:

$$\left\{\left[b_1, F^{-1}_{X_1|A}(t)\right] \times \cdots \times \left[b_n, F^{-1}_{X_n|A}(t)\right] \mid t \in [0, 1]\right\}.$$

From our assumption that $f$ is continuous and increasing in each $X_i$,

$$\begin{aligned} &f\left(\left[b_1, F^{-1}_{X_1|A}(t)\right] \times \cdots \times \left[b_n, F^{-1}_{X_n|A}(t)\right]\right) \\ &\quad = \left[f(b_1, \ldots, b_n), f\left(F^{-1}_{X_1|A}(t), \ldots, F^{-1}_{X_n|A}(t)\right)\right]. \end{aligned}$$

Thus,

$$F_{Y|A}\left(f\left(F^{-1}_{X_1|A}(t), \ldots, F^{-1}_{X_n|A}(t)\right)\right) \geq F_{X|A}\left(F^{-1}_{X_1|A}(t), \ldots, F^{-1}_{X_n|A}(t)\right).$$

Because for a level curve $c = y = f(\vec{x})$, $\vec{x} \in [\underline{s}_1, \bar{s}_1] \times \cdots \times [\underline{s}_n, \bar{s}_n]$, the assumption that $f$ is increasing in each variable means that for a fixed $\vec{x}^*$ such that $c = f(\vec{x}^*)$, the level curve lies in $[\underline{s}_1, x_1^*] \times \cdots \times [\underline{s}_i, x_i^*] \times \cdots \times [\underline{s}_n, x_n^*]$ for each $i = 1, \ldots, n$. This can be seen in Figure 7 for the example $Y = f(X_1, X_2) = X_1X_2$ with the level curve $c = \frac{1}{10} = f(X_1, X_2) = X_1X_2$ so that $\vec{x}^* = \{(x_1, x_2) \mid |x_1x_2 = \frac{1}{10}\}$. If $X_1$ and $X_2$ are i.i.d. $U[0, 1]$ random variables with $Y = f(X_1, X_2) = X_1X_2$, and $\vec{x}^* = (\sqrt{\frac{1}{10}}, \sqrt{\frac{1}{10}})$, then clearly $f(\vec{x}^*) \in [0, \sqrt{\frac{1}{10}}] \times [0, \sqrt{\frac{1}{10}}]$ since the right end point of this box is a point on the level curve.

Now consider the $n$-dimensional box (coming from the right end point) $[F^{-1}_{X_1|A}(t), c_1] \times \cdots \times [F^{-1}_{X_n|A}(t), c_n]$. As before, we know that

$$\begin{aligned} &f\left(\left[F^{-1}_{X_1|A}(t), c_1\right] \times \cdots \times \left[F^{-1}_{X_n|A}(t), c_n\right]\right) \\ &\quad = \left[f\left(F^{-1}_{X_1|A}(t), \ldots, F^{-1}_{X_n|A}(t)\right), f(c_1, \ldots, c_n)\right]. \end{aligned}$$

This gives the inequality

$$1 - F_{Y|A}\left(f\left(F^{-1}_{X_1|A}(t), \ldots, F^{-1}_{X_n|A}(t)\right)\right) \geq G_{X|A}\left(F^{-1}_{X_1|A}(t), \ldots, F^{-1}_{X_n|A}(t)\right).$$

Combining both inequalities yields

$$\begin{aligned} &F_{X|A}\left(F^{-1}_{X_1|A}(t), \ldots, F^{-1}_{X_n|A}(t)\right) \\ &\quad \leq F_{Y|A}\left(f\left(F^{-1}_{X_1|A}(t), \ldots, F^{-1}_{X_n|A}(t)\right)\right) \\ &\quad \leq 1 - G_{X|A}\left(F^{-1}_{X_1|A}(t), \ldots, F^{-1}_{X_n|A}(t)\right). \end{aligned} \tag{77}$$

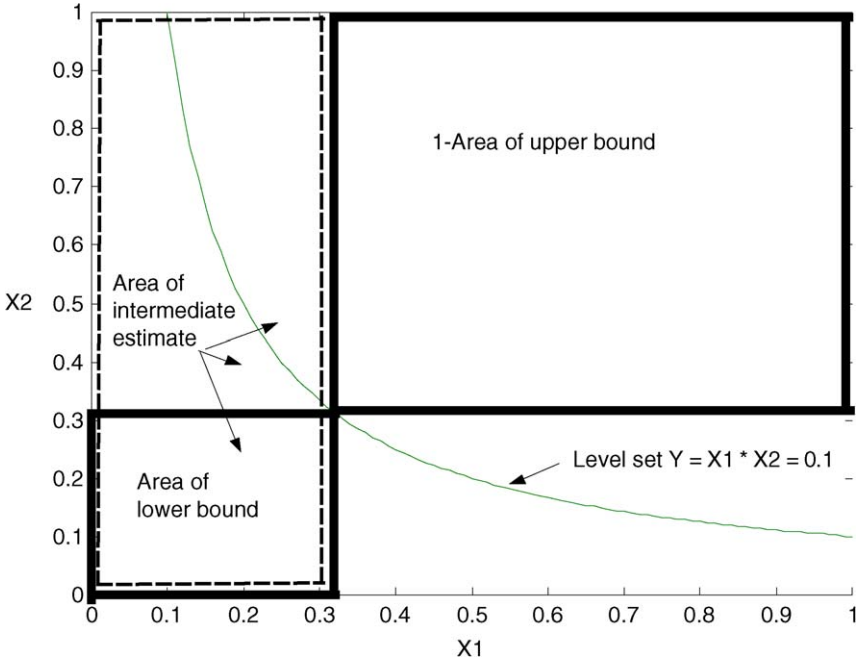


FIGURE 7. Outer and inner approximations on $Y = 0.1$ isoline. Outer and inner measure of $F(Y) = 0.1$, $Y$-product of two i.i.d. $U[0, 1]$ random variables with no partitioning (one block).

***Example.*** Figure 7 illustrates the estimate and bounds for $F_Y(0.1)$ when $Y = X_1X_2$ and $X_i$ are identical independent uniform [0, 1] random variables. The range has not been subdivided. This example shows the inner and outer measures of the area bounded to the left and right of the level set $Y = X_1X_2 = 0.1$, the lower and upper box computed for $\vec{x}^* = (\sqrt{\frac{1}{10}}, \sqrt{\frac{1}{10}})$, illustrating the inequalities of Eq. (77). NOTE: the two lower and upper bounds can be computed from any $\vec{x}^*$ such that $\vec{x}^* = \{(x_1, x_2) \mid x_1x_2 = \frac{1}{10}\}$. The algorithm that is developed here uses points along the diagonal to make the size of the boxes the same (hypersquares).

A simplification to this calculation is accomplished by the use of two auxiliary functions, $g$ and $h$, defined below. To this end, define

$$g : [0, 1]^n \to \prod_{i=1}^{n} [\underline{s}_i, \bar{s}_i] \tag{78}$$

by

$$g(u_1, \ldots, u_n) = \left(F_{X_1}^{-1}(u_1), \ldots, F_{X_n}^{-1}(u_n)\right) = (x_1, \ldots, x_n). \tag{79}$$

By the assumptions, $g$ is onto, strictly increasing in each variable, thus one-to-one. Thus, it is invertible so that

$$g(U) = \vec{X} = (X_1, \ldots, X_n)$$

and

$$U = g^{-1}(X_1, \ldots, X_n).$$

It is clear that from Eqs. (78) and (79) that $U_i = F_{X_i}(X_i)$ [from which we derive $X_i = F_{X_i}^{-1}(U_i)$] and has the standard uniform distribution

$$U_i(t) = \begin{cases} 1/(\bar{s}_i - \underline{s}_i) & \text{for } t \in [\underline{s}_i, \bar{s}_i] \\ 0 & \text{otherwise.} \end{cases} \tag{80}$$

This is because $U_i$ takes values in the interval $[0, 1]$, and the probability that $U_i$ is less than or equal to $t$ is the probability that $F_{X_i}(X_i) \leq t$. But this is the probability that $X_i$ is in the interval $(-\infty, c]$, where $c$ is the constant such that $F_{X_i}(c) = t$. Thus, the probability that $X_i$ is in the interval $(-\infty, c]$ is precisely $t$, which means that $U_i$ is the standard uniform distribution [Eq. (80)]. In general, if a random variable is the domain of its cumulative distribution function, a standard uniform distribution is obtained.

Fix $y^* \in [\underline{s}_Y, \bar{s}_Y]$ and consider the set $A = \{(x_1, \ldots, x_n) \mid f(x_1, \ldots, x_n) \leq y^*\}$, which would be the area under the isoline in the example shown in Figure 7. Then

$$F_Y(y^*) = prob(X \in A) = prob\left(U \in g^{-1}(A)\right), \tag{81}$$

where Eq. (81) is equal to the area of the set

$$\left\{(u_1, \ldots, u_n) \mid g(u_1, \ldots, u_n) \in A\right\} = \left\{(u_1, \ldots, u_n) \mid f\left(g(u_1, \ldots, u_n)\right) \leq y^*\right\}.$$

This means that the problem of finding the cumulative distribution function of Eq. (58) can be redefined as a problem of finding the cumulative distribution function for the random variable

$$Y = h(U) = f\left(g(U)\right), \tag{82}$$

where

$$h : [0, 1]^n \to [\underline{s}_Y, \bar{s}_Y],$$
$$h(u_1, \ldots, u_n) = f\left(F_{X_1}^{-1}(u_1), \ldots, F_{X_n}^{-1}(u_n)\right),$$

where $U_i$ is Eq. (80). This is because $Y$, defined by Eq. (82), has the identical distribution as $Y = f(X)$. Next define

$$k : [0, 1] \to [\underline{s}_Y, \bar{s}_Y] \tag{83}$$

by

$$k(t) = h(t, \ldots, t), \quad t \in [0, 1]. \tag{84}$$

By assumption, $k$ is one-to-one, onto, strictly increasing and invertible so that $\exists t^*$ such that $t^* = k^{-1}(y^*)$. Let $T = k^{-1}(Y)$. Then,

$$F_T\left(t^*\right) = F_Y\left(y^*\right), \tag{85}$$

which is the area of the set

$$\left\{(u_1, \ldots, u_n) \mid h(u_1, \ldots, u_n) \le y^*\right\}$$
$$= \left\{(u_1, \ldots, u_n) \mid h(u_1, \ldots, u_n) \le h\left(t^*, \ldots, t^*\right)\right\}.$$

Given that $h$ is monotonically increasing in each variable, for $u_i \le t^* \ \forall i$, we know

$$h(u_1, \ldots, u_n) \le h\left(t^*, \ldots, t^*\right),$$

so that this area is bounded below by $(t^*)^n$. Similarly, for $u_i \ge t^* \ \forall i$,

$$h(u_1, \ldots, u_n) \ge h\left(t^*, \ldots, t^*\right),$$

so that this area is bounded above $1 - (1 - t^*)^n$. This means that the following bounds hold:

$$\left(t^*\right)^n \le F_T\left(t^*\right) = F_Y\left(y^*\right) \le 1 - \left(1 - t^*\right)^n. \tag{86}$$

Of course, an open question is how to pick bounds that are a simple computation such as that given above.

The bounds [Eq. (86)] may be a wide envelope, particularly for a large number of variables (large $n$). A reasonable estimate of the cumulative distribution function, without having to perform the calculations needed to reduce the envelope to a small width, is desired. To do this, select an intermediate value $\widehat{F}_{Y|A}$ for

$$F_{Y|A}\left(f\left(F_{X_1|A}^{-1}(t), \ldots, F_{X_n|A}^{-1}(t)\right)\right)$$

that falls between the lower and upper estimate above. One estimate would be to average these probabilities, that is, set

$$\begin{aligned}\widehat{F}_{Y|A}\big(f\big(F^{-1}_{X_1|A}(\beta t),\ldots,F^{-1}_{X_n|A}(t)\big)\big)&\\ =\frac{1}{2}\big(F_{X|A}\big(F^{-1}_{X_1|A}(t),\ldots,F^{-1}_{X_n|A}(t)\big)&\\ +1-G_{X|A}\big(F^{-1}_{X_1|A}(t),\ldots,F^{-1}_{X_n|A}(t)\big)\big).&\end{aligned}$$

When the random variables, $\vec{X}_i$, are independent, a reasonable choice for the intermediate estimate function, $\widehat{F}$, is to use $t$. This works since $t^n \le t \le 1-(1-t)^n$ and has several desirable properties. First is its simplicity. Second is that this estimate does not increase the maximum possible error in making a choice of intermediate value. This is so because the maximum of the difference $1-(1-t)^n-t^n$ occurs when $t=0.5$, and at this value the midpoint estimate is $\frac{1}{2}(0.5^n+1-0.5^n)=0.5$. A third property is that it is symmetric about the value $t=0.5$. Any tendency to overestimate or underestimate the true value when $t<0.5$ should be offset by a tendency to underestimate or overestimate for values of $t>0.5$. So, for independent $\vec{X}$ we use

$$\underline{F}_{Y|A}\big(f\big(F^{-1}_{X_1|A}(t),\ldots,F^{-1}_{X_n|A}(t)\big)\big)=t^n, \tag{87}$$

$$\widehat{F}_{Y|A}\big(f\big(F^{-1}_{X_1|A}(t),\ldots,F^{-1}_{X_n|A}(t)\big)\big)=t, \tag{88}$$

and

$$\overline{F}_{Y|A}\big(F^{-1}_{X_1|A}(t),\ldots,F^{-1}_{X_n|A}(t)\big)=1-(1-t)^n \tag{89}$$

to obtain a lower bound, intermediate estimate, and upper bound, respectively, on the actual cumulative distribution function, $F_{Y|A}(y)$. Note that these equations [Eqs. (87)–(89)], define three functions of $t$, $H_k(t):[0,1]\to[0,1]$, $k=1,2,3$. When $n=3$, the graphs of $H_k$ for each of the three estimates are as follows (see Figure 8).

The actual distribution lies between the lower and upper bounds, while the estimate $H_2(t)=t$ provides a centrally located estimate. This is the source of the averaging of errors as the number of subdivisions increases. As the number of variables becomes large, the lower and upper bounds may become wide and the intermediate estimate becomes more important.

*Step 3. Combine*— The last step of the method is to combine the estimated conditional cumulative distribution functions (CDF) into an approximation of the cumulative distribution function for $Y$. Assume the support has been divided into subintervals for each $X_i$ creating a partition of the support of $\vec{X}$. If $\{A_j \mid j=1,m\}$ is the partition (so each $A_j$ is also an $n$-dimensional box), and if $\underline{F}_{Y|A_i}(y)$, $\widehat{F}_{Y|A_i}(y)$, and $\overline{F}_{Y|A_i}(y)$ are the values calculated as

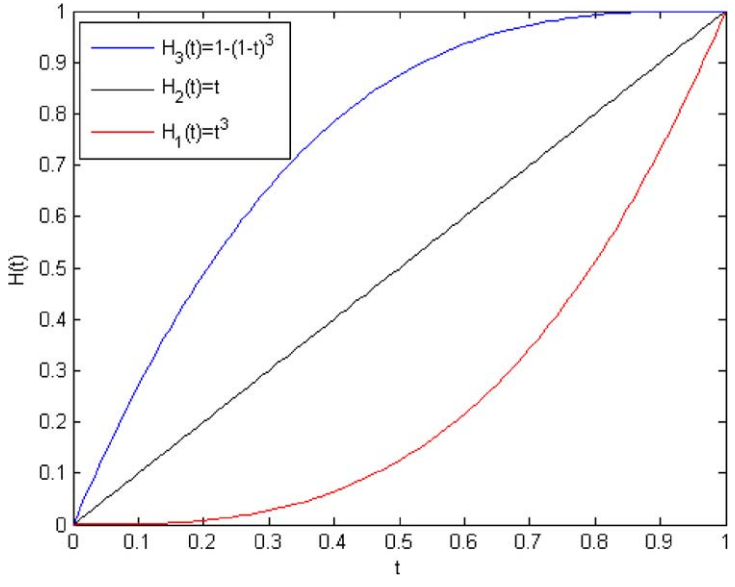


FIGURE 8. Underestimate, approximation, and overestimate.

above for the lower bound, intermediate approximation, and the upper bound (respectively) of the cumulative distribution function for the variable $Y$ given that $X \in A_i$, then we can combine these cumulative distribution functions to produce the bounds and estimate for the cumulative distribution function of interest by setting $\widehat{F}_Y(y) = \sum_{j=1}^{m} \widehat{F}_{Y|A_i}(y) P(X \in A_j)$, where $P(E)$ equals the probability of the event $E$. The lower and upper bounds $[\underline{F}_Y(y)$ and $\overline{F}_Y(y)]$ are calculated similarly.

***Example.*** Consider $Y = f(X_1, X_2) = X_1 + X_2$ with 2 partitions (4 boxes), the same problem with 8 partitions (64 boxes) and a more complicated function $Y = (\max\{X_1^3, X_2\})^2 + X_1 X_2 X_3$ with 8 partitions (512 boxes), where $X_1$, $X_2$, and $X_3$ are identical, independent uniform over [0, 1] (i.i.d. $\tilde{}U[0, 1]$) illustrated in Figures 10 and 11, respectively. Note that while the lower and upper bounds are wide for the sum of two random variables using four boxes (Figure 9), the estimate is very good. If an estimate suffices, then computing with fewer boxes works well. All simulated results, using Monte Carlo used 10,000 random draws.

*Theoretical Considerations.* The algorithm above converges to $F_Y(y)$ when the supports of $X_i$ are subdivided in such a way that the diameters of all

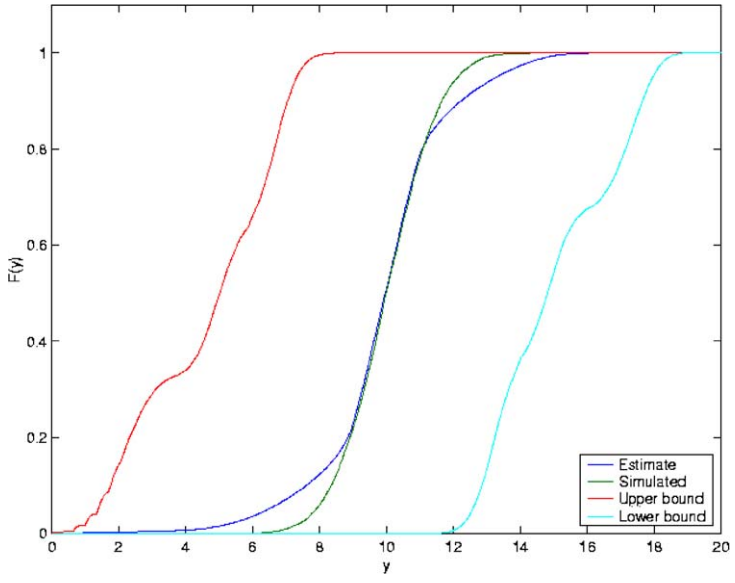


FIGURE 9. Sum of 20 i.i.d. $U[0, 1]$ random variables using 4 boxes, $2 \times 2$ partitions.

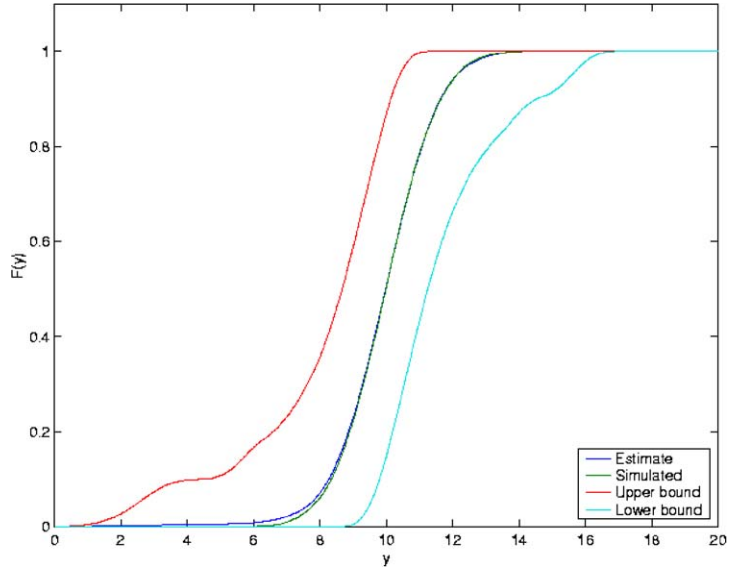


FIGURE 10. Sum of 20 i.i.d. $U[0, 1]$ random variables using 64 boxes, $8 \times 8$ partitions.

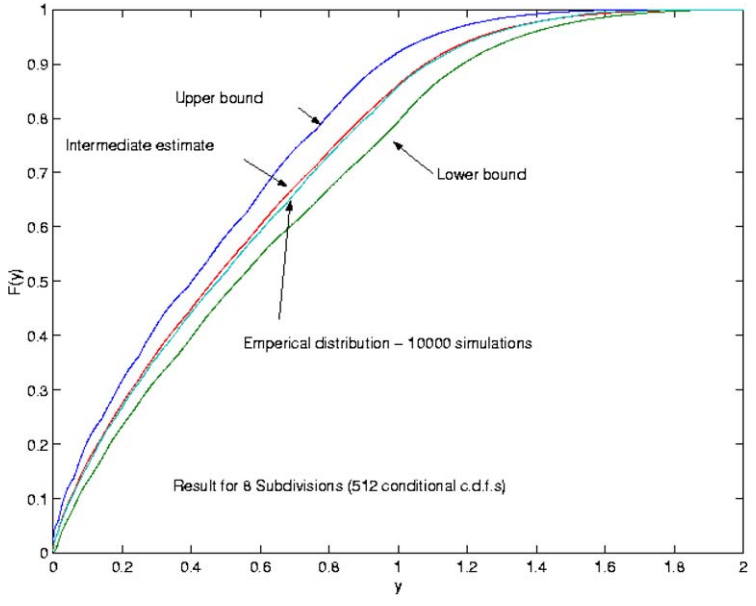


FIGURE 11. Distribution arithmetic using Jamison and Lodwick on a complex problem using 512 boxes, $8 \times 8 \times 8$ partitions. The cumulative distribution function for $Y = (\max\{X_1^3, X_2\})^2 + X_1, X_2, X_3$ when $X_1$, $X_2$, and $X_3$ are i.i.d. $U[0, 1]$ random variables.

subboxes go to 0, that is, when

$$diam(D_k) = diam\left(\left[\underline{s}_1^k, \bar{s}_1^k\right] \times \cdots \times \left[\underline{s}_n^k, \bar{s}_n^k\right]\right) \to 0 \quad \forall k.$$

This is because, from measure theory and our assumptions, $\underline{F_Y}(y)$ and $\overline{F_Y}(y)$ are simply inner and outer measures for $F_Y$ relative to the measure defined by $F_Y$ for the region $\{\vec{x} \mid f(\vec{x}) \leq y\}$. Thus, the process will converge to the actual cumulative distribution function.

**b. Olsen.** Olsen (2005) particularizes Jamison and Lodwick with the view of obtaining an efficient algorithm suitable for computing the function of random variables in high dimensions. Olsen uses Tonon's approach to construct the resulting cumulative distribution function. Thus, this approach may be considered a hybrid of Jamison and Lodwick and Tonon with the intended purpose of obtaining an efficient algorithm that encloses the actual distribution. The theoretical considerations are those of the Jamison and Lodwick method discussed above. The Olsen algorithm is as follows.

***Algorithm IV.5*** (Olsen, 2005). **Find the cumulative distribution of** $y = f(X_1, \ldots, X_n)$, where the $X_i$ are *independent random variables*.

*Step 1.* Partition $X_i$ into $m$ intervals of equal probability, $p$, as was done in the Jamison and Lodwick method, resulting in

$$X_i = \left\{X_i^1, \ldots, X_i^m\right\}, \quad \text{where } X_i^j = \left[\underline{x}_i^j, \bar{x}_i^j\right],$$

so that

$$\begin{aligned} A_1 &= X_1^1 \times X_2^1 \times \cdots \times X_{n-1}^1 \times X_n^1, \\ A_2 &= X_1^1 \times X_2^1 \times \cdots \times X_{n-1}^1 \times X_n^2, \\ &\vdots \\ A_{n^m} &= X_1^m \times X_2^m \times \cdots \times X_{n-1}^m \times X_n^m, \end{aligned}$$

and probability of each $A_k$ is $p^n$ given our independence assumption.

*Step 2.* Compute the minimum and maximum of $f(A_j)$. NOTE: It is here that what Tonon (2004) and Fortin *et al.* (2006) call the *vertex method*, is used. The vertex method is simply what is known in optimization—the minimum and maximum of a monotone function over a compact set (box) occurs at the end points. Olsen (2005) assumes that the function is monotone over each box and thus simply evaluates $f(A_j)$ at the vertices of boxes.

*Step 3.* Construct the lower and upper bounds for the cumulative distribution function like Tonon (2004), that is,

$$\underline{F}(y) = \sum_{y \geq \sup f(A_i)} p(A_i) = p^n \sum_{y \geq \sup f(A_i)}, \tag{90}$$

$$\overline{F}(y) = \sum_{y > \inf(A_i)} p(A_i) = p^n \sum_{y > \inf f(A_i)}. \tag{91}$$

Since the range of $f$ is an interval $[\underline{y}, \bar{y}]$, one typically picks a set of grid points $y_k \in [\underline{y}, \bar{y}] y_0 = \underline{y} < y_1 \cdots < y_k < y_{k+1} < \cdots < y_K = \bar{y}$, where Eqs. (90) and (91) are evaluated at the points of the grid selected.

***Example.*** Let $Y = f(X_1, X_2) = X_1 + X_2$, for $X_1, X_2$ $U[0, 1]$, and independent.

*Step 1. Partition* $X_i$ into $m$ intervals of equal probability, say, $p = 0.5$. Thus, $X_{11} = [0, \frac{1}{2}]$, $X_{12} = [\frac{1}{2}, 1]$, $X_{21} = [0, \frac{1}{2}]$, $X_{22} = [\frac{1}{2}, 1]$, so that, $A_1 = [0, \frac{1}{2}] \times [0, \frac{1}{2}]$, $A_2 = [0, \frac{1}{2}] \times [\frac{1}{2}, 1]$, $A_3 = [\frac{1}{2}, 1] \times [0, \frac{1}{2}]$, $A_4 = [\frac{1}{2}, 1] \times [\frac{1}{2}, 1]$, each having probability of $p = (\frac{1}{2})^2 = \frac{1}{4}$, since we have assumed independence. This is illustrated in Figure 12.

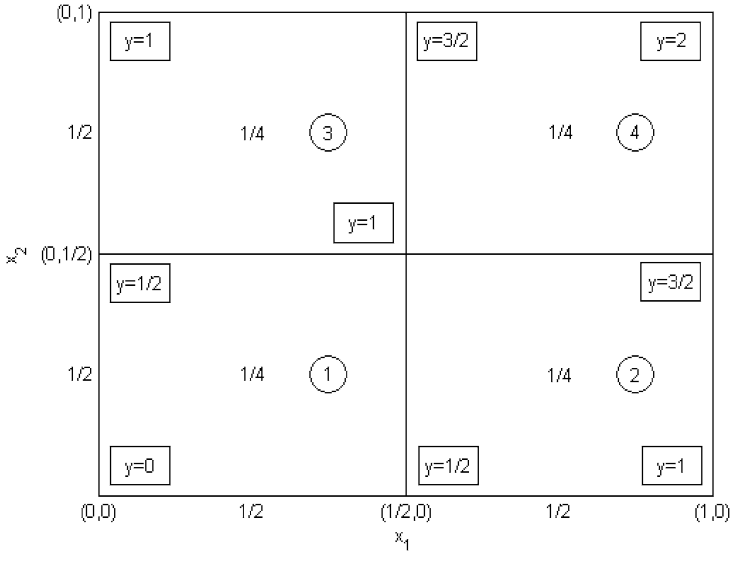


FIGURE 12. Example of inverse probability boxes with equal probability.

*Step 2. Compute the minimum and maximum* of $f(X_1, X_2)$, which occurs at the vertices $\{(0,0), (0,\frac{1}{2}), (0,1), (\frac{1}{2},0), (\frac{1}{2},\frac{1}{2}), (\frac{1}{2},1), (1,0), (1,\frac{1}{2}), (1,1)\}$. Since the order does not matter for addition, the minimum and maximum of $f$ on the vertices will be $(0,0) = 0+0 = 0, (0,\frac{1}{2}) = 0+\frac{1}{2} = 0.5, \ldots (1,1) = 1+1 = 2$. The function values of $f$ on each of the four $A_j$s at the vertices are $B_1 = \{0, \frac{1}{2}, \frac{1}{2}, 1\}$, $B_2 = \{\frac{1}{2}, \frac{1}{2}, 1, \frac{3}{2}\}$, $B_3 = \{\frac{1}{2}, \frac{1}{2}, 1, \frac{3}{2}\}$, and $B_4 = \{1, \frac{3}{2}, \frac{3}{2}, 2\}$ with the minimum and maximum on $B_1$ being $\{0, 1\}$; on $B_2$ it is $\{\frac{1}{2}, \frac{3}{2}\}$; on $B_3$ it is $\{\frac{1}{2}, \frac{3}{2}\}$; and on $B_4$ it is $\{1, 2\}$.

*Step 3. Construct the lower and upper bounds* on the cumulative distribution by taking the minimum and maximum of each $B_j$. Thus for the grid $y_0 = 0$, $y_1 = \frac{1}{2}$, $y_2 = 1$, $y_3 = \frac{3}{2}$, $y_4 = 2$, begin with the smallest value $y_0 = 0$. This value is not greater than any of the $B_j$ values, so that $\overline{F}(0) = 0$. At $y_1 = \frac{1}{2}$, $y_1$ is greater than the minimum of $B_1$ but no other $B_j$s, so $\overline{F}(\frac{1}{2}) = \frac{1}{4}$, the probability of one $A_j$. Continuing the process, we obtain $y_i < \min\{B_j\}\overline{F}(1) = \frac{3}{4}$, $\overline{F}(\frac{3}{2}) = 1$, $\overline{F}(2) = 1$. Similarly, for $y_i \geq \max\{B_j\}\underline{F}(0) = 0$, $\underline{F}(\frac{1}{2}) = 0$, $\underline{F}(1) = 0$, $\underline{F}(\frac{3}{2}) = \frac{3}{4}$, $\underline{F}(2) = 1$. The lower and upper cumulative distributions and the Moore approximation are illustrated in Figure 13. (See also Tables 2 and 3.)

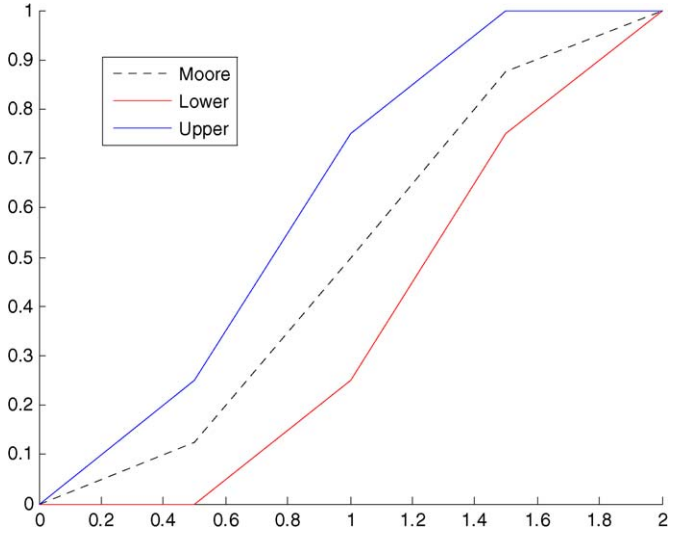


FIGURE 13. Inverse probability of Olsen—assembled lower and upper cumulative distribution function.

TABLE 2

| $y_k > \min\{B_j\} \Rightarrow p(B_j) \in$ c.d.f. $F_{y_k}$ | $B_1$ | $B_2$ | $B_3$ | $B_4$ | CDF |
|---|---|---|---|---|---|
| $y_0 = 0$ | 0 | 0 | 0 | 0 | 0 |
| $y_1 = \frac{1}{2}$ | $\frac{1}{4}$ | 0 | 0 | 0 | $\frac{1}{4}$ |
| $y_2 = 1$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | 0 | $\frac{3}{4}$ |
| $y_3 = \frac{3}{2}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | 1 |
| $y_4 = 2$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | 1 |

***Remark.*** Functions that are nonmonotonic in a box, for the methods that evaluate over boxes, require global optima to be computed. When dependencies are unknown, global optimization over boxes is the only way to obtain a meaningful result when enclosures are necessary. From lower and upper approximations, a guess at the actual distribution from this information is often useful. The examples have used averages that for independent variables are reasonable. There are two areas that would help in obtaining a more efficient method for the computation of enclosure of resulting distribution from a function of distributions.

1. Strategies for partitioning that are efficient in terms of obtaining global optima *and* computing the probability over boxes. Global optimization

TABLE 3

| $y_k \le \max\{B_j\} \Rightarrow p(B_j) \in$ c.d.f. $F_{y_k}$ | $B_1$ | $B_2$ | $B_3$ | $B_4$ | CDF |
|---|---|---|---|---|---|
| $y_0 = 0$ | 0 | 0 | 0 | 0 | 0 |
| $y_1 = \frac{1}{2}$ | 0 | 0 | 0 | 0 | 0 |
| $y_2 = 1$ | $\frac{1}{4}$ | 0 | 0 | 0 | $\frac{1}{4}$ |
| $y_3 = \frac{3}{2}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | 0 | $\frac{3}{4}$ |
| $y_4 = 2$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | 1 |

methods that provide enclosure and verification require interval techniques. This means that partitioning strategies are useful (see Hansen, 1992; Kearfott, 1996a; Neumaier, 2004b). In particular, second derivative information, if available, may be useful in deciding the partitioning strategy.

2. As seen from the examples, the lower and upper approximations that enclose the distribution provide the bounds from which a guess to the actual distribution can be made. The examples use the average between the lower and upper approximation, which for independent variables is a reasonable approach. However, good strategies to pick a "best" approximation from the lower and upper bounds are needed. There is a trade-off between the number of partitions and the resulting complexity. One way to reduce the complexity is to have a good approximation method based on the information provided by the lower and upper bounds when the number of partitions is relatively low.

### *B. General Theory of Uncertainty*

A problem in which some or all the types of uncertainty of interest (interval, probability, possibility, and fuzzy) occur within the problem requires a theory that has the uncertainties of interest as special instances of the theoretical framework. Two theories do this: (1) the theory of clouds (Neumaier, 2004a), and (2) the theory of interval-valued probability (Jamison and Lodwick, 2006; Lodwick and Jamison, 2006; Weichselberger, 1996, 2000), which can be considered as extensions of Walley (1991, 1999). Since interval probability measures include clouds (as well as intervals, probability, possibility and fuzzy uncertainty), the more detailed focus will be on interval-valued probability measures. This discussion begins with the theory of clouds.

#### 1. *Clouds*

The idea of a cloud is to enclose uncertainty in such a way that the enclosing functions have probabilistic-like characteristics. In particular, every cloud has

been shown to contain a probability distribution within it (Neumaier, 2005). Beyond the ability to model with a mixture of uncertainty, the original impetus was to be able to model analytically when missing information, precision of concepts, models, and/or measurements.

***Definition 25.*** A *cloud* (Neumaier, 2004a) over a set $M$ is a mapping $\mathbf{x}$ that associates with each $\xi \in M$ a (nonempty, closed, and bounded interval) $\mathbf{x}(\xi)$, such that,

$$(0, 1) \subseteq \bigcup_{\xi \in M} \mathbf{x}(\xi) \subseteq [0, 1]. \tag{92}$$

$\mathbf{x}(\xi) = [\underline{\mathbf{x}}(\xi), \bar{\mathbf{x}}(\xi)]$, is called the *level* of $\xi$ in the cloud $\mathbf{x}$, where $\underline{\mathbf{x}}(\xi)$ and $\bar{\mathbf{x}}(\xi)$ are the *lower* and *upper level*, respectively, and $\bar{\mathbf{x}}(\xi) - \underline{\mathbf{x}}(\xi)$ is called the *width* of $\xi$. When the width is 0 for all $\xi$, the cloud is called a *thin cloud*.

When doing analysis over real numbers, a concept akin to a fuzzy number (gradual number or interval number in our previous settings) is required for clouds.

***Definition 26*** (Neumaier, 2004a). A real *cloudy number* is a cloud over the set $\mathbb{R}$ of real numbers. $\chi_{[a,b]}$ ($\chi$ being the characteristic function) is the cloud equivalent to an interval $[a, b]$, providing support information without additional probabilistic content. A *cloudy vector* is cloud over $\mathbb{R}^n$, where each component is a cloudy number.

Neumaier (2004a) states that dependence or correlation between uncertain numbers (or the lack thereof) can be modeled by considering them jointly as components of a cloudy vector. Moreover,

> In many applications (not always, cf. Proposition IV.1, but roughly when $\bar{\mathbf{x}}(\xi) \approx 1$ for $\xi$ near the modes of the associated distribution), the level $\mathbf{x}(\xi)$ may be interpreted as giving lower and upper bounds on the *degree of suitability* of $\xi \in M$ as a possible scenario for data modeled by the cloud $\mathbf{x}$. This degree of suitability can be given a probability interpretation by relating clouds to random variables [see Eq. (93) below].

We say that a random variable $x$ with values in $M$ belongs to a cloud $\mathbf{x}$ over $M$, and write $x \in \mathbf{x}$, if

$$\Pr\big(\underline{x}(x) \geq \alpha\big) \leq 1 - \alpha \leq \Pr\big(\bar{x}(x) > \alpha\big), \quad \forall \alpha \in [0, 1]. \tag{93}$$

Pr denotes the probability of the statement given as argument, and it is required that the sets consisting of all $\xi \in M$ where $\underline{x}(x) \geq \alpha$ and $\bar{x}(x) > \alpha$

are measurable in the $\sigma$-*algebra* on $M$ consisting of all sets $A \subseteq M$ for which $\Pr(x \in A)$ is defined. This approach gives clouds an underlying interpretation as the class of random variables $x$ with $x \in \mathbf{x}$. A fuzzy cloud is one for which the level is a pair of fuzzy membership functions. The above interpretation is equivalent to the interpretation of fuzzy set membership degree as an upper bound (for a single fuzzy set considered as a cloud, have the upper level as the membership function, and the $x$-axis as the lower level) for probabilities first advocated by Dubois *et al.* (1997). Since there exists at least one random variable in every cloud (Neumaier, 2005), this interpretation is meaningful.

Computation with clouds involves arithmetic on the (vertical) interval $\mathbf{x}(\zeta)$ for each $\xi$ that defines the lower and upper level of the cloud. Efficient methods to compute with clouds are still an open question. However, once a cloud is constructed, methods that have been discussed can be used.

***Example*** (Kawai, 2006). Suppose a histogram-based random variable $x$ is given in Table 4.

The original density is not crisp. The density of Figure 14 assumes uniform distribution within each interval. If the $X_i$ are chosen in the order shown in Figure 14, then we have a discrete cloud that resembles Figure 15.

The dashed lines of Figure 15 form $\underline{\mathbf{x}}(\xi)$ (the lower level) and the solid line of Figure 15 form $\bar{\mathbf{x}}(\xi)$ (the upper level). There is nothing to prevent the cloud from being formed using a different order of the $X_i$. In fact, there may be some advantage in choosing the intervals in a different order so that it results in a bell-shaped cloud, as shown in Figure 16.

## 2. *Interval-Valued Probability Measures*

The next presentation is an expanded version of Lodwick and Jamison (2006) and portions may also be found in Jamison and Lodwick (2006). A basis for

TABLE 4

| Intervals $X_i$ | | $\Pr(x \in X_i)$ | Walley (1999) uses a cumulative $\alpha_i$ | $\mathbf{x}(\xi) := [\alpha_{i-1}, \alpha_i]$ |
|---|---|---|---|---|
| $X_1$ | $(-\infty, 2)$ | 0.0 | 0.0 | [0.0, 0.0] |
| $X_2$ | [2, 3) | 0.1 | 0.1 | [0.0, 0.1] |
| $X_3$ | [3, 4) | 0.2 | 0.3 | [0.1, 0.3] |
| $X_4$ | [4, 5) | 0.3 | 0.6 | [0.3, 0.6] |
| $X_5$ | [5, 6] | 0.4 | 1.0 | [0.6, 1.0] |
| $X_6$ | $(6, \infty)$ | 0.0 | 1.0 | [1.0, 1.0] |

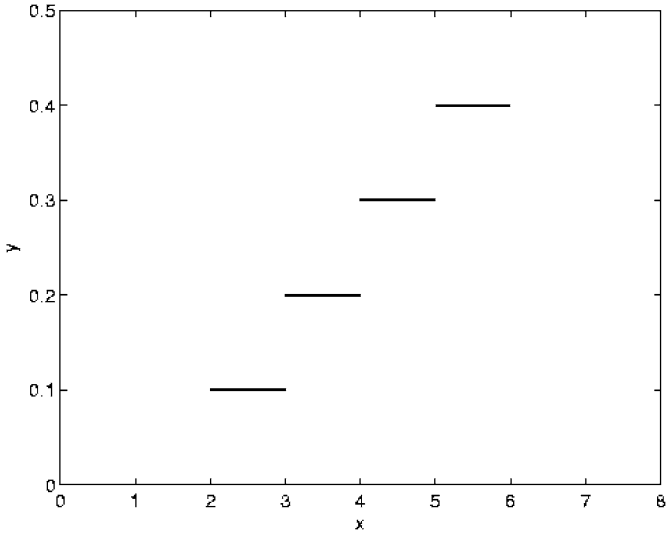


FIGURE 14. Probability density—histogram-based random variable.

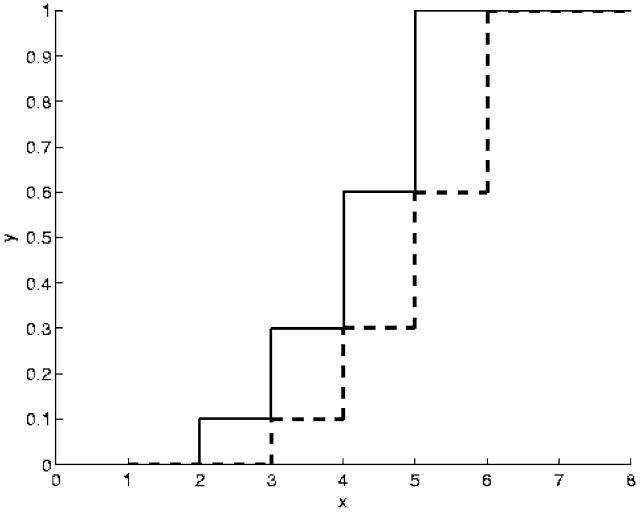


FIGURE 15. Discrete cloud constructed from probability density in order given.

linking various methods of uncertainty representation, including clouds, is examined next. This section begins by defining what is meant by an IVPM. This generalization of a probability measure includes probability measures,

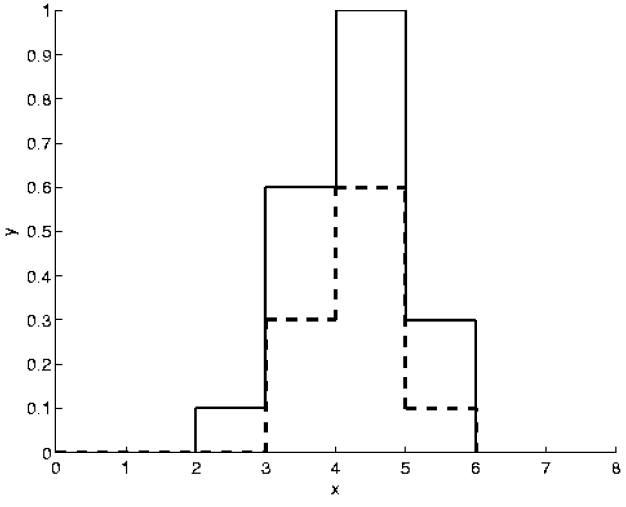


FIGURE 16. Discrete cloud constructed from density in different order than that given.

possibility and necessity measures, intervals, and clouds (Neumaier, 2004a). The set function defining IVPM is thought of as a method for providing a partial representation for an unknown probability measure, much like clouds. Throughout, arithmetic operations involving set functions are in terms of interval arithmetic (Moore, 1979), and $Int_{[0,1]} \equiv \{[a, b] \mid 0 \leq a \leq b \leq 1\}$ denotes an arbitrary interval within [0, 1]. It will be seen how problems using mixed representations can be handled and solved.

The first subsection defines in a formal way the term *interval-valued probability* measure as used by Weichselberger (1996, 2000). Weichselberger's definition begins with a set of probability measures (our IVPMs relax "tightest") and then defines an interval probability as a set function providing lower and upper bounds on the probabilities calculated from these measures. *F-probabilities* are simply the tightest bounds possible for the set of probability measures. This definition is followed by demonstrating that various forms of uncertainty representation (possibility, interval, cloud, and probability) all can be represented by such measures. The next subsection shows how IVPMs can be constructed from lower and upper bounding cumulative distribution functions. This is followed by an extension principle for a function of uncertain variables represented by IVPMs and integration with respect to IVPMs. Both definitions will be useful in analyzing problems involving uncertainty represented by IVPMs. An application to a problem in optimization is given.

Throughout this section, we are interested primarily in interval-valued probability defined on the Borel sets on the real line and real-valued random variables. The basic definitions from Weichselberger (with slight variation in notation) are presented next.

***Definition 27*** (Weichselberger, 2000). Given measurable space $(S, \mathcal{A})$, an interval-valued function $i_m : \mathcal{A} \to Int_{[0,1]}$ is called an *R-probability* if:

(a) $i_m(A) = [a^-(A), a^+(A)] \subseteq [0, 1]$ with $a^-(A) \leq a^+(A)$
(b) There exists a probability measure Pr on $\mathcal{A}$ such that

$$\forall A \in \mathcal{A}, \quad \Pr(A) \in i_m(A).$$

By an *R-probability field*, we mean the triple $(S, \mathcal{A}, i_m)$.

***Definition 28*** (Weichselberger, 2000). Given an $R$-probability field $\mathcal{R} = (S, \mathcal{A}, i_m)$ the set

$$\mathcal{M}(\mathcal{R}) = \{\Pr \mid \Pr \text{ is a probability measure on } \mathcal{A} \text{ such that } \forall A \in \mathcal{A}, \Pr(A) \in i_m(A)\}$$

is called the *structure* of $\mathcal{R}$.

***Definition 29*** (Weichselberger, 2000). An $R$-probability field $\mathcal{R} = (S, \mathcal{A}, i_m)$ is called an *F-probability field*, if $\forall A \in \mathcal{A}$:

(a) $a^+(A) = \sup\{\Pr(A) \mid \Pr \in \mathcal{M}(\mathcal{R})\}$,
(b) $a^-(A) = \inf\{\Pr(A) \mid \Pr \in \mathcal{M}(\mathcal{R})\}$.

It is interesting to note that given a measurable space $(S, \mathcal{A})$ and a set of probability measures $P$, then defining $a^+(A) = \sup\{\Pr(A) \mid \Pr \in P\}$ and $a^-(A) = \inf\{\Pr(A) \mid \Pr \in P\}$ gives an *F-probability*, where $P$ is a subset of the structure.

The following examples show how intervals, possibility distributions, clouds, and (of course) probability measures can define *R-probability* fields on $\mathcal{B}$, the Borel sets on the real line.

***Example*** (An interval defines an $F$-probability field). Let $I = [a, b]$ be a nonempty interval on the real line. On the Borel sets, define

$$a^+(A) = \begin{cases} 1 & \text{if } I \cap A \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

and

$$a^-(A) = \begin{cases} 1 & \text{if } I \subseteq A \\ 0 & \text{otherwise,} \end{cases}$$

then $i_m(A) = [a^-(A), a^+(A)]$ defines an $F$-probability field $\mathcal{R} = (R, \mathcal{B}, i_m)$. To see this, simply set $P$ to be the set of all probability measures on $\mathcal{B}$ such that $\Pr(I) = 1$.

***Example*** (A probability measure is an $F$-probability field). Let Pr be a probability measure over $(S, \mathcal{A})$. Define $i_m(A) = [\Pr(A), \Pr(A)]$, which is equivalent to having total knowledge about a probability distribution over $S$.

The concept of a cloud was introduced by Neumaier (2004a), and in the context of the notation of this section, it is defined as follows.

***Definition 30.*** A cloud over set $S$ is a mapping $c$ such that:

(1) $\forall s \in S, c(s) = [\underline{n}(s), \bar{p}(s)]$ with $0 \leq \underline{n}(s) \leq \bar{p}(s) \leq 1$,
(2) $(0, 1) \subseteq \bigcup_{s \in S} c(s) \subseteq [0, 1]$.

In addition, random variable $X$ taking values in $S$ is said to belong to cloud $c$ (written $X \in c$) if

(3) $\forall \alpha \in [0, 1], \Pr(\underline{n}(X) \geq \alpha) \leq 1 - \alpha \leq \Pr(\bar{p}(X) > \alpha)$.

Clouds are closely related to possibility theory. A function $p : S \to [0, 1]$ is called a regular possibility distribution function if

$$\sup\{p(x) \mid x \in S\} = 1.$$

Possibility distribution functions (Wang and Klir, 1992) define a possibility measure,

$$Pos : S \to [0, 1],$$

where

$$Pos(A) = \sup\{p(x) \mid x \in A\},$$

and its dual necessity measure,

$$Nec(A) = 1 - Pos(A^c).$$

By convention, we define $\sup\{p(x) \mid x \in \emptyset\} = 0$. A necessity distribution function can also be defined as

$$n : S \to [0, 1]$$

by setting

$$n(x) = 1 - p(x).$$

Observe that

$$Nec(A) = \inf\{n(x) \mid x \in A^c\},$$

where we define $\inf\{n(x) \mid x \in \emptyset\} = 1$. Jamison and Lodwick (2002) showed that possibility distributions could be constructed that satisfy the following consistency definition.

***Definition 31.*** Let $p : S \to [0, 1]$ be a regular possibility distribution function with associated possibility measure *Pos* and necessity measure *Nec*. Then $p$ is said to be *consistent* with random variable $X$ if for all measurable sets $A$, $Nec(A) \leq \Pr(X \in A) \leq Pos(A)$.

***Remark.*** Recall that a distribution acts on real numbers, and measures act on sets of real numbers.

The concept of a cloud can be stated in terms of certain pairs of consistent possibility distributions, which is shown in the following proposition.

**Proposition IV.1.** *Let $\bar{p}$, $\underline{p}$ be a pair of regular possibility distribution functions over set $S$ such that $\forall s \in S\ \bar{p}(s) + \underline{p}(s) \geq 1$. Then the mapping $c(s) = [\underline{n}(s), \bar{p}(s)]$, where $\underline{n}(s) = 1 - \underline{p}(s)$ (that is, the dual necessity distribution function) is a cloud. In addition, if $X$ is a random variable taking values in $S$ and the possibility measures associated with $\bar{p}$, $\underline{p}$ are consistent with $X$, then $X$ belongs to cloud $c$. Conversely, every cloud defines such a pair of possibility distribution functions, and their associated possibility measures are consistent with every random variable belonging to the cloud $c$.*

*Proof.*

$\Rightarrow$

(1) $\bar{p}, \underline{p} : S \to [0, 1]$ and $\bar{p}(s) + \underline{p}(s) \geq 1$ imply property (1) of Definition 31.
(2) Since all regular possibility distributions satisfy $\sup\{p(s) \mid s \in S\} = 1$ property (2) of Definition 31 holds.

Therefore $c$ is a cloud. Now assume consistency. Then $\alpha \geq Pos\{s \mid p(s) \leq \alpha\} \geq \Pr\{s \mid p(s) \leq \alpha\} = 1 - \Pr\{s \mid p(s) > \alpha\}$ gives the right-hand side of the required inequalities and

$$\begin{aligned} 1 - \alpha &\geq Pos\{s \mid p(s) \leq 1 - \alpha\} \\ &\geq \Pr\{s \mid p(s) \leq 1 - \alpha\} \\ &= \Pr\{s \mid 1 - p(s) \geq \alpha\} \\ &= \Pr(\underline{n}(X) \geq \alpha) \end{aligned}$$

gives the left-hand side.

$\Leftarrow$ The opposite identity was proven in Section 5 of Neumaier (2004a). □

***Example*** (A cloud defines an $R$-probability field). Let $c$ be a cloud over the real line. If $Pos^1, Nec^1, Pos^2, Nec^2$ are the possibility measures and their dual necessity measures relating to $\bar{p}(s)$ and $\underline{p}(s)$, define

$$i_m(A) = \big[\max\{Nec^1(A), Nec^2(A)\}, \min\{Pos^1(A), Pos^2(A)\}\big].$$

Neumaier (2005) proved that every cloud contains a random variable $X$. Consistency requires that $\Pr(X \in A) \in i_m(A)$ and thus every cloud defines an $R$-probability field.

***Example*** (A possibility distribution defines an $R$-probability field). Let $p : S \to [0, 1]$ be a regular possibility distribution function, and let $Pos$ be the associated possibility measure and $Nec$ the dual necessity measure. Define $i_m(A) = [Nec(A), (A)]$. If we define a second possibility distribution, $\underline{p}(x) = 1\ \forall x$, then the pair $p$, $\underline{p}$ define a cloud for which $i_m(A)$ defines the $R$-probability.

**a. Construction from Kolmogorov–Smirnoff Statistics.** In this subsection, an *F-probability* is constructed from lower and upper bounding cumulative distribution functions in a manner allowing practical computation. For example, given statistical data, we can construct a confidence interval for the underlying cumulative distribution function using the Kolmogorov estimation of confidence limits (see Kolmogorov, 1941). Then using this confidence interval, we can use the following development to construct an IVPM. Although a simple definition could be used by simply setting the interval equal to the lower and upper bound over all probability measures contained in the bound, it is not clear how to use this definition in practice. The development that follows is more amenable to actual use.

Let

$$F^u(x) = \Pr\big(X^u \le x\big),$$

and

$$F^l(x) = \Pr\big(X^l \le x\big),$$

be two cumulative distribution functions for random variables over the Borel sets on the real line, $X^u$ and $X^l$, with the property that $F^u(x) \ge F^l(x)\ \forall x$. Set

$$\mathcal{M}\big(X^u, X^l\big) = \big\{X \mid \forall x\ F^u(x) \ge \Pr(X \le x) \ge F^l(x)\big\},$$

which clearly contains $X^u$ and $X^l$. We will think in terms of an unknown $X \in \mathcal{M}(X^u, X^l)$. For any Borel set $A$, let $\Pr(A) = \Pr(X \in A)$.

We begin by developing probability bounds for members of the family of sets

$$\mathcal{I} = \{(a, b], (-\infty, a], (a, \infty), (-\infty, \infty), \emptyset \mid a < b\}.$$

For $I = (-\infty, b]$, it is clear by definition that

$$\Pr(I) \in [F^l(b), F^u(b)].$$

For $I = (a, \infty)$, let

$$\Pr(I) \in [1 - F^u(a), 1 - F^l(a)].$$

For $I = (a, b]$, since $I = (-\infty, b] - (-\infty, a]$, and considering minimum and maximum probabilities in each set, let

$$\Pr(I) \in [\max\{F^l(b) - F^u(a), 0\}, F^u(b) - F^l(a)].$$

Therefore, if we extend the definition of $F^u$, and $F^l$ by defining

$$F^u(-\infty) = F^l(-\infty) = 0,$$

and

$$F^u(\infty) = F^l(\infty) = 1,$$

we can make the following general definition.

***Definition 32.*** For any $I \in \mathcal{I}$, if $I \neq \emptyset$, define

$$i_m(I) = [a^-(I), a^+(I)] = [\max\{F^l(b) - F^u(a), 0\}, F^u(b) - F^l(a)],$$

where $a$ and $b$ are the left and right end points of $I$. Otherwise, set

$$i_m(\emptyset) = [0, 0].$$

***Remark.*** Note that with this definition

$$\begin{aligned} i_m((-\infty, \infty)) &= [\max\{F^l(\infty) - F^u(-\infty), 0\}, F^u(\infty) - F^l(-\infty)] \\ &= [1, 1], \end{aligned}$$

which matches our intuition and thus, it is easy to see that $\Pr(I) \in i_m(I)$ $\forall I \in \mathcal{I}$.

We can extend this to include finite unions of elements of $\mathcal{I}$. For example, if

$$E = I_1 \cup I_2 = (a, b] \cup (c, d]$$

with $b < c$, then we consider the probabilities

$$\Pr\big((a, b]\big) + \Pr\big((c, d]\big),$$
$$1 - \big(\Pr\big((-\infty, a]\big) + \Pr\big((b, c]\big) + \Pr\big((d, \infty)\big)\big)$$

(the probability of the sets that make up $E$ versus one less the probability of the intervals that make up the complement), and consider the minimum and maximum probability for each case as a function of the minimum and maximum of each set. The *minimum* for the first sum is

$$\max\big(0, F^l(d) - F^u(c)\big) + \max\big(0, F^l(b) - F^u(a)\big),$$

and the *maximum* is

$$F^u(d) - F^l(c) + F^u(b) - F^l(a).$$

The minimum for the second is

$$\begin{aligned} &1 - \big(F^u(\infty) - F^l(d) + F^u(c) - F^l(b) + F^u(a) - F^l(-\infty)\big) \\ &\quad = F^l(d) - F^u(c) + F^l(b) - F^u(a), \end{aligned}$$

and the maximum is

$$\begin{aligned} &1 - \big(\max\big(0, F^l(\infty) - F^u(d)\big) \\ &\qquad + \max\big(0, F^l(c) - F^u(b)\big) + \max\big(0, F^l(a) - F^u(-\infty)\big)\big) \\ &\quad = F^u(d) - \max\big(0, F^l(c) - F^u(b)\big) - F^l(a). \end{aligned}$$

This gives

$$\Pr(E) \geq \max \begin{cases} F^l(d) - F^u(c) + F^l(b) - F^u(a) \\ \max\big(0, F^l(d) - F^u(c)\big) + \max\big(0, F^l(b) - F^u(a)\big) \end{cases}$$

and

$$\Pr(E) \leq \min \begin{cases} F^u(d) - \max\big(0, F^l(c) - F^u(b)\big) - F^l(a) \\ F^u(d) - F^l(c) + F^u(b) - F^l(a) \end{cases},$$

so

$$\begin{aligned} \Pr(E) \in \big[&\max\big(0, F^l(d) - F^u(c)\big) + \max\big(0, F^l(b) - F^u(a)\big), \\ &F^u(d) - \max\big(0, F^l(c) - F^u(b)\big) - F^l(a)\big]. \end{aligned}$$

The final line is arrived at by noting that

$$\forall x, \quad y F^l(x) - F^u(y) \leq \max\big(0, F^l(x) - F^u(y)\big).$$

***Remark.*** Note the two extreme cases for $E = (a, b] \cup (c, d]$. For $F^u(x) = F^l(x) = F(x)\ \forall x$, then, as expected,

$$\Pr(E) = F(d) - F(c) + F(b) - F(a) = \Pr\big((a, b]\big) + \Pr\big((c, d]\big),$$

that is, it is the probability measure. Moreover, for $F^l(x) = 0\ \forall x$,

$$\Pr(E) \in \big[0,\ F^u(d)\big],$$

that is, it is a possibility measure for the possibility distribution function $F^u(x)$.

Let

$$\mathcal{E} = \left\{ \bigcup_{k=1}^{K} I_k \;\middle|\; I_k \in \mathcal{I} \right\}.$$

That is, $\mathcal{E}$ is the algebra of sets generated by $I$. Note that every element of $E$ has a unique representation as a union of the minimum number of elements of $\mathcal{I}$ (or, stated differently, as a union of disconnected elements of $\mathcal{I}$). Note also that $R \in \mathcal{E}$ and $\mathcal{E}$ is closed under complements.

Assume $E = \bigcup_{k=1}^{K} I_k$ and $E^c = \bigcup_{j=1}^{J} M_j$ are the unique representations of $E$ and $E^c$ in $\mathcal{E}$ in terms of elements of $\mathcal{I}$. Then, considering minimum and maximum possible probabilities of each interval, it is clear that

$$\Pr(E) \in \left[ \max\left( \sum_{k=1}^{K} a^-(I_k),\ 1 - \sum_{j=1}^{J} a^+(M_j) \right), \right.$$
$$\left. \min\left( \sum_{k=1}^{K} a^+(I_k),\ 1 - \sum_{j=1}^{J} a^-(M_j) \right) \right].$$

This can be made more concise using the following result.

**Proposition IV.2.** *If $E = \bigcup_{k=1}^{K} I_k$ and $E^c = \bigcup_{j=1}^{J} M_j$ are the unique representations of $E$ and $E^c \in \mathcal{E}$, then $\sum_{k=1}^{K} a^-(I_k) \geq 1 - \sum_{j=1}^{J} a^+(M_j)$, and $\sum_{k=1}^{K} a^+(I_k) \geq 1 - \sum_{j=1}^{J} a^-(M_j)$.*

*Proof.* We need only prove

$$\sum_{k=1}^{K} a^-(I_k) \geq 1 - \sum_{j=1}^{J} a^+(M_j),$$

since we can exchange the roles of $E$ and $E^c$, giving

$$\sum_{j=1}^{J} a^-(M_j) \geq 1 - \sum_{k=1}^{K} a^+(I_k),$$

thereby proving the second inequality. Note $\sum_{k=1}^{K} a^-(I_k) + \sum_{j=1}^{J} a^+(M_j)$ is of the form

$$\sum_{k=1}^{K} \max\bigl(0, F^l(b_k) - F^u(a_k)\bigr) + \sum_{j=1}^{J} F^u(a_{j+1}) - F^l(b_j)$$
$$\geq \sum_{k=1}^{K} \bigl(F^l(b_k) - F^u(a_k)\bigr) + \sum_{j=1}^{J} F^u(a_{j+1}) - F^l(b_j).$$

Since the union of the disjoint intervals yields all of the real line, we have either $F^u(\infty)$ or $F^l(\infty)$ less either $F^u(-\infty)$ or $F^l(-\infty)$, which is 1 regardless. □

Next $i_m$ is extended to $\mathcal{E}$.

**Proposition IV.3.** *For any $E \in \mathcal{E}$, let $E = \bigcup_{k=1}^{K} I_k$, and $E^c = \bigcup_{j=1}^{J} M_j$ be the unique representations of $E$ and $E^c$ in terms of elements of $\mathcal{I}$, respectively. If*

$$i_m(E) = \left[\sum_{k=1}^{K} a^-(I_k), 1 - \sum_{j=1}^{J} a^-(M_j)\right],$$

*then $i_m : \mathcal{E} \to Int_{[0,1]}$ is an extension of $\mathcal{I}$ to $\mathcal{E}$ and is well defined. In addition,*

$$i_m(E) = \bigl[\inf\bigl\{\Pr(X) \in E \mid X \in \mathcal{M}\bigl(X^u, X^l\bigr)\bigr\},$$
$$\sup\bigl\{\Pr(X) \in E \mid X \in \mathcal{M}\bigl(X^u, X^l\bigr)\bigr\}\bigr].$$

*Proof.* First assume $E = (a, b] \in \mathcal{I}$, then $E^c = (-\infty, a] \cup (b, \infty)$, so by the definition,

$$i_m(E) = \bigl[\max\bigl\{F^l(b) - F^u(a), 0\bigr\},$$
$$1 - \bigl(\max\bigl\{F^l(a) - F^u(-\infty), 0\bigr\} + \max\bigl\{F^l(\infty) - F^u(b), 0\bigr\}\bigr)\bigr],$$

which matches the definition for $i_m$ on $\mathcal{I}$. The other cases for $E \in \mathcal{I}$ are similar. Thus, it is an extension. It is easy to show that it is well defined, since the representation of any element in $\mathcal{E}$ in terms of the minimum number of

elements of $\mathcal{I}$ is unique. In addition, it is clear that

$$0 \leq \sum_{k=1}^{K} a^{-}(I_k)$$

and

$$1 - \sum_{j=1}^{J} a^{-}(M_j) \leq 1.$$

So we only need to show that

$$\sum_{k=1}^{K} a^{-}(I_k) \leq 1 - \sum_{j=1}^{J} a^{-}(M_j).$$

That is,

$$\sum_{k=1}^{K} a^{-}(I_k) + \sum_{j=1}^{J} a^{-}(M_j) \leq 1.$$

If we relabel the end points of all these intervals as $-\infty = c_1 < c_2 < \cdots < c_N = \infty$, then

$$\begin{aligned}
\sum_{k=1}^{K} a^{-}(I_k) + \sum_{j=1}^{J} a^{-}(M_j) &= \sum_{n=1}^{N-1} \max\left\{F^{l}(c_{n+1}) - F^{u}(c_n), 0\right\} \\
&\leq \sum_{n=1}^{N-1} \max\left\{F^{u}(c_{n+1}) - F^{u}(c_n), 0\right\} \\
&= \sum_{n=1}^{N-1} F^{u}(c_{n+1}) - F^{u}(c_n) \\
&= 1.
\end{aligned}$$

Thus

$$\sum_{k=1}^{K} a^{-}(I_k) + \sum_{j=1}^{J} a^{-}(M_j) \leq 1.$$

For the last equation, assume

$$E = \bigcup_{k=1}^{K} I_k = (-\infty, b_1] \cup (a_2, b_2] \cup \cdots \cup (a_K, b_K]$$

and

$$E^c = \bigcup_{j=1}^{J} M_j = (b_1, a_2] \cup \cdots \cup (b_K, \infty).$$

We will show that

$$X \in \mathcal{M}\big(X^u, X^l\big) \quad \Rightarrow \quad \Pr(X \in E) \in i_m(E),$$

and there is an $X \in \mathcal{M}(X^u, X^l)$ for which $\Pr(X \in E) = a^+(E)$. Note first that

$$\begin{aligned}\sum_{j=1}^{J} a^-(M_j) &= \sum_{k=1}^{K} \max\big\{F^l(a_{k+1}) - F^u(b_k), 0\big\} \\ &\le \sum_{k=1}^{K} \max\big\{F(a_{k+1}) - F(b_k), 0\big\} \\ &= \Pr\big(E^c\big),\end{aligned}$$

which gives both

$$\Pr(E) = 1 - \Pr\big(E^c\big) \le a^+(E),$$

and by replacing $E$ with

$$E^c a^-(E) \le \Pr(E).$$

Next for $x \le a_2$, set

$$F(x) = \min\big(F^l(b_1),\, F^u(x)\big),$$

and for $a_2 < x \le b_2$, set

$$F(x) = \min\left(F^l(b_2),\, \left(\frac{x - a_2}{b_2 - a_2}\right)F^u(x) + \left(\frac{b_2 - x}{b_2 - a_2}\right)F^u(x)\right).$$

Continuing in this way gives a cumulative distribution function for which

$$\Pr\big(E^c\big) = \sum_{j=1}^{J} a^-(M_j)$$

and

$$\Pr(E) = 1 - \sum_{j=1}^{J} a^-(M_j).$$

The other bound is similarly derived. □

The family of sets, $\mathcal{E}$, is a ring of sets generating the Borel sets $\mathcal{B}$. For an arbitrary Borel set $S$, then it is clear that

$$\Pr(S) \in \left[\sup\left\{a^-(E) \mid E \subseteq S, E \in \mathcal{E}\right\}, \inf\left\{a^+(F) \mid S \subseteq F, F \in \mathcal{E}\right\}\right].$$

This prompts the following:

**Proposition IV.4.** *Let $i_m : \mathcal{B} \to [0, 1]$ be defined by*

$$i_m(A) = \left[\sup\left\{a^-(E) \mid E \subseteq A, E \in \mathcal{E}\right\}, \inf\left\{a^+(F) \mid A \subseteq F, F \in \mathcal{E}\right\}\right].$$

*Then $i_m$ is an extension from $\mathcal{E}$ to $\mathcal{B}$, and it is well defined.*

*Proof.* The last property of Proposition IV.3 ensures it is an extension since, for example, if $E \subseteq F$ are two elements of $\mathcal{E}$, then $a^+(E) \leq a^+(F)$ so

$$\inf\left\{a^+(F) \mid E \subseteq F, F \in \mathcal{E}\right\} = a^+(E).$$

Similarly, it ensures that

$$\sup\left\{a^-(F) \mid F \subseteq E, F \in \mathcal{E}\right\} = a^-(E).$$

Next we show that $i_m$ is well defined. Proposition IV.3 shows that

$$\forall E \in \mathcal{E}, \quad i_m(E) \subseteq [0, 1].$$

Thus,

$$0 \leq \sup\left\{a^-(E) \mid E \subseteq S\right\}$$

and

$$\inf\left\{a^+(E) \mid S \subseteq E\right\} \leq 1.$$

We also have

$$\sup\left\{a^-(E) \mid E \subseteq S\right\} \leq \inf\left\{a^+(F) \mid S \subseteq F\right\}.$$

□

**Proposition IV.5.** *The function $i_m : \mathcal{B} \to Int_{[0,1]}$ defines an F-probability field on the Borel sets and*

$$i_m(B) = \left[\inf\left\{\Pr(X \in B) \mid X \in \mathcal{M}\left(X^u, X^l\right)\right\}, \sup\left\{\Pr(X \in B) \mid X \in \mathcal{M}\left(X^u, X^l\right)\right\}\right],$$

*that is, $\mathcal{M}(X^u, X^l)$ defines a structure.*

*Proof.* Clear. □

**b. Interval-Valued Integration, Extension and Independence of *F*-Probabilities.** This subsection defines three key concepts needed for the application of IVPMs to mathematical programming problems: integration, extension, and independence.

***Definition 33.*** Given $F$-probability field $\mathcal{R} = (S, \mathcal{A}, i_m)$ and an integrable function $f : S \to R$, we define:

$$\int_A f(x)\, di_m = \left[ \inf_{p \in \mathcal{M}(\mathcal{R})} \int_A f(x)\, dp, \sup_{p \in \mathcal{M}(\mathcal{R})} \int_A f(x)\, dp \right].$$

We make the following observations that are useful in actual evaluation.

It is easy to see that if $f$ is an $\mathcal{A}$-measurable simple function such that

$$f(x) = \begin{cases} y & x \in A \\ 0 & x \notin A \end{cases} \quad \text{with } A \in \mathcal{A},$$

then

$$\int_A f(x)\, di_m = y i_m(A).$$

Further, if $f$ is a simple function taking values $\{y_k \mid k \in K\}$ on an at-most countable collection of disjoint measurable sets $\{A_k \mid k \in K\}$ that is,

$$f(x) = \begin{cases} y_k & x \in A_k \\ 0 & x \notin A \end{cases} \quad \text{where } A = \bigcup_{k \in K} A_k,$$

then

$$\int_A f(x)\, di_m = \left[ a^- \left( \int_A f(x)\, di_m \right), a^+ \left( \int_A f(x)\, di_m \right) \right],$$

where

$$a^+ \left( \int_A f(x)\, di_m \right) = \sup \left\{ \sum_{k \in K} y_k \Pr(A_k) \;\middle|\; \Pr \in \mathcal{M}(\mathcal{R}) \right\}$$

and

$$a^- \left( \int_A f(x)\, di_m \right) = \inf \left\{ \sum_{k \in K} y_k \Pr(A_k) \;\middle|\; \Pr \in \mathcal{M}(\mathcal{R}) \right\}. \tag{94}$$

Note that these can be evaluated by solving two linear programming problems, since $\Pr \in \mathcal{M}(\mathcal{R})$ implies that $\sum_{k \in K} \Pr(A) = 1$, and $\Pr(\bigcup_{l \in L \subset K} A_l) \in i_m(\bigcup_{l \in L \subset K} A_l)$ so the problem may be tractable. In general, if $f$ is an

integrable function, and $\{f_k\}$ is a sequence of simple functions converging uniformly to $f$, then the integral with respect to $f$ can be determined by noting that

$$\int_A f(x)\,di_m = \lim_{k\to\infty} \int_A f_k(x)\,di_m,$$

where

$$\lim_{k\to\infty} \int_A f_k(x)\,di_m = \left[\lim_{k\to\infty} a^-\left(\int_A f_k(x)\,di_m\right), \lim_{k\to\infty} a^+\left(\int_A f_k(x)\,di_m\right)\right],$$

provided the limits exist.

***Example.*** Consider the IVPM constructed from the interval $[a, b]$. Then $\int_R x\,di_m = [a, b]$, that is, the interval-valued expected value is the interval itself.

***Definition 34.*** Let $\mathcal{R} = (S, \mathcal{A}, i_m)$ be an $F$-probability field and $f : S \to T$ a measurable function from measurable space $(S, \mathcal{A})$ to measurable space $(T, \mathcal{B})$. Then the $F$-probability $(T, \mathcal{B}, l_m)$ defined by

$$l_m(B) = \left[\inf\left\{\Pr\left(f^{-1}(B)\right) \mid \Pr \in \mathcal{M}(\mathcal{R})\right\}, \sup\left\{\Pr\left(f^{-1}(B)\right) \mid \Pr \in \mathcal{M}(\mathcal{R})\right\}\right] \tag{95}$$

is called the extension of the $R$-probability field to $(T, \mathcal{B})$.

That this defines an *F-probability* field is clear from our earlier observation. In addition, it is easy to see that this definition is equivalent to setting

$$l_m(A) = i_m\left(f^{-1}(A)\right),$$

which allows for evaluation using the techniques described earlier.

We now address the combination of IVPMs when the variables are independent. We do not address the situation when dependencies may be involved. Given measurable spaces $(S, \mathcal{A})$ and $(T, \mathcal{B})$ and the product space $(S \times T, \mathcal{A} \times \mathcal{B})$, assume $i_{X\times Y}$ is an IVPM on $\mathcal{A} \times \mathcal{B}$. Call $i_X$ and $i_Y$ defined by $i_X(A) = i_{X\times Y}(A \times T)$ and $i_Y(B) = i_{X\times Y}(S \times B)$ the marginals of $i_{X\times Y}$. The marginals, $i_X$ and $i_Y$, are IVPMs.

***Definition 35.*** Call the marginal IVPMs independent if and only if $i_{X\times Y}(A \times B) = i_X(A) i_Y(B)\ \forall A, B \subseteq S$.

***Definition 36.*** Let $\mathcal{R} = (S, \mathcal{A}, i_X)$ and $\mathcal{Q} = (T, \mathcal{B}, l_Y)$ be $F$-probability fields representing uncertain random variables $X$ and $Y$. We define the $F$-probability field $(S \times T, \mathcal{A} \times \mathcal{B}, i_{X\times Y})$ by defining

$$i^+_{X\times Y}(A \times B) = \sup\Big\{\Pr_X(B)\Pr_Y(B) \;\Big|\; \Pr_X \in \mathcal{M}(\mathcal{R}), \Pr_Y \in \mathcal{M}(\mathcal{Q})\Big\},$$

$$i^-_{X\times Y}(A \times B) = \inf\Big\{\Pr_X(B)\Pr_Y(B) \;\Big|\; \Pr_X \in \mathcal{M}(\mathcal{R}), \Pr_Y \in \mathcal{M}(\mathcal{Q})\Big\},$$

where $(S \times T, \mathcal{A} \times \mathcal{B})$ is the usual product of sigma algebra of sets.

It is clear from this definition that

$$i_{X\times Y}(A \times B) \equiv i_X(A) i_Y(B)$$

for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Thus, if we have several uncertain parameters in a problem with the uncertainty characterized by IVPMs, and all are independent, we can form an IVPM for the product space by multiplication and use this IVPM.

**c. Application to Optimization.** An example application of IVPM is an optimization problem, the recourse problem. Suppose we wish to optimize $f(x, a)$ subject to $g(x, b) = 0$. Assume $a$ and $b$ are vectors of independent uncertain parameters, each with an associated IVPM. Assume the constraint can be violated at a known cost $c$ so that the problem is to solve:

$$h(x, a, b) = \max\big(f(x, a) - c\big(g(x, b)\big)\big).$$

Using the independence, form an IVPM for the product space, $i_{a\times b}$, for the joint distribution. Then calculate the interval-valued expected value with respect to this IVPM. The resulting interval-valued expected value is

$$\int_R h(x, a, b)\, di_{a\times b}.$$

To optimize over such a value requires an ordering of intervals. One such ordering is to use the midpoint of the interval on the principle that in the absence of additional data, the midpoint is the best estimate for the true value. Another possible ordering is to use risk/return multi-objective decision making. For example, determine functions $u : \mathbb{R}^2 \to \mathbb{R}$ and $v : Int_{\mathbb{R}} \to \mathbb{R}^2$ by setting, for any interval $I = [a, b]$, $v(I) = (\frac{a+b}{2}, b - a)$. Thus, $v$ gives the midpoint and width of an interval. Then $u$ would measure the decision makes preference for one interval over another considering both its midpoint

and width (a risk measure). The optimization problem becomes

$$\max_x u\left(v\left(\int_R h(x,a,b)\,di_{a\times b}\right)\right).$$

### C. Generalized Extension Principles for Distribution

One extension principle associated with general distributions is presented above [Eq. (95)], and is perhaps the only one that is able to deal with the complete set of uncertainty that is of interest to this article. The current research seems to lack a generalized extension principle except what is presented by Jamison and Lodwick (2006) and that given here. The reason for this is not only the complexity of the problem, but a lack of a general theory that captures the broad spectrum of uncertainty distribution, which IVPMs of Weichselberger (2000) and Jamison and Lodwick (2006) do.

## REFERENCES

Aberth, O. (1978). A method for exact computation with rational numbers. *Journal of Computational and Applied Mathematics* **4**, 285–288.

Aberth, O. (1988). *Precise Numerical Analysis*. William C. Brown, Dubuque, IA.

Alefeld, G., Herzberger, J. (1983). *Introduction to Interval Computations*. Academic Press, New York.

Alefeld, G. (1990). Enclosure methods. In: Ullrich, C. (Ed.), *Computer Arithmetic and Self-Validating Numerical Methods*. Academic Press, Boston, MA, pp. 55–72.

Anile, M., Deodato, S., Privitera, G. (1995). Implementing fuzzy arithmetic. *Fuzzy Sets and Systems* **72** (2), 239–250.

Archimedes of Siracusa (1897). Measurement of a circle. In: Heath, T.L. (Ed.), *The Works of Archimedes*. Cambridge University Press, Cambridge. Dover edition, 1953.

Audin, J.-P., Frankkowska, H. (1990). *Set-Valued Analysis*. Birkhäuser, Boston, MA.

Baudrit, C., Dubois, D., Fargier H. (2005). Propagation of uncertainty involving imprecision and randomness. ISIPTA, 31–40.

Berleant, D. (1993). Automatically verified reasoning with both intervals and probability density functions. *Interval Computations* **2**, 48–70.

Berleant, D., Goodman-Strauss, C. (1998). Bounding the results of arithmetic operations on random variables of unknown dependency using intervals. *Reliable Computing* **4**, 147–165.

Berleant, D., Zhang, J. (2004). Representation and problem solving with distribution envelope determination (DEnv). *Reliability Engineering and Systems Safety* **85** (1–3), 153–168.

Böhmer, K., Hemker, P., Stetter, H.J. (1984). The defect correction approach. *Computing Supplement* **5**, 1–32.

Bondia, J., Sala, A., Sáinz, M. (2006). Modal fuzzy quantities and applications to control. In: Demirli, K., Akgunduz, A. (Eds.), *2006 Conference of the North American Fuzzy Information Processing Society*, June 3–6, 2006, Montréal, Canada.

Bornemann, F., Laurie, D., Wagon, S., Waldvogel, J. (2004). *The SIAM 100-Digit Challenge: A Study in High-Accuracy Numerical Computing*. SIAM, Philadelphia, PA.

Burkill, J.C. (1924). Functions of intervals. *Proceedings of the London Mathematical Society* **22**, 375–446.

Corliss, G.F. (2004). Tutorial on validated scientific computing using interval analysis. PARA'04 Workshop on State-of-the-Art Computing. Technical University of Denmark, June 20–23, 2004. See http://www.eng.mu.edu/corlissg/PARA04/READ_ME.html.

Daney, D., Papegay, Y., Neumaier, A. (2004). Interval methods for certification of the kinematic calibration of parallel robots. In: *Proc. 2004 IEEE Int. Conf. Robotics Automation*. New Orleans, LA, pp. 191–198.

Davies, B. (2005). Whither mathematics? *Notices of the AMS* **52** (11), 1350–1356.

Demster, A.P. (1967). Upper and lower probabilities induced by multivalued mapping. *Annals of Mathematical Statistics* **38**, 325–339.

Dempster, M.A.H. (1969). Distributions in interval and linear programming. In: Hansen, E.R. (Ed.), *Topics in Interval Analysis*. Oxford University Press, pp. 107–127.

Dempster, M.A.H. (1974). An application of quantile arithmetic to the distribution problem in stochastic linear programming. *Bulletin of the Institute of Mathematics and Its Applications* **10**, 186–194.

Dubois, D., Prade, H. (1977). Le flou, mécédonka, Tech. Rep. C.E.R.T.-D.E.R.A. Toulouse, April 1977.

Dubois, D., Prade, H. (1978). Operations on fuzzy numbers. *International Journal of Systems Science* **9** (6), 613–626.

Dubois, D., Prade, H. (1980). *Fuzzy Sets and Systems: Theory and Applications*. Academic Press, New York.

Dubois, D., Prade, H. (1981). Additions of interactive fuzzy numbers. *IEEE Trans. on Automatic Control* **26** (4), 926–936.

Dubois, D., Prade, H. (1987a). Evidence theory and interval analysis. *Second IFSA Congress, Tokyo, July 20–25, 1987*, pp. 502–505.

Dubois, D., Prade, H. (1987b). Fuzzy numbers: An overview. Technical Report 219, L.S.I., Univ. Paul Sabatier, Toulouse, France; also in Bezdek, J.C. (Ed.), *Analysis of Fuzzy Information: Mathematics and Logic, vol. 1*. CRC Press, Boca Raton, FL, pp. 3–39.

Dubois, D., Prade, H. (1988). *Possibility Theory—An Approach to Computerized Processing of Uncertainty*. Plenum Press, New York.

Dubois, D., Prade, H. (1991). Random sets and fuzzy interval analysis. *Fuzzy Sets and Systems* **42**, 87–101.

Dubois, D., Prade, H. (Eds.) (2000a). *Fundamentals of Fuzzy Sets*. Kluwer Academic Press.

Dubois, D., Prade, H. (2005). Fuzzy elements in a fuzzy set. In: *Proceedings of the 10th International Fuzzy System Association (IFSA) Congress*, Beijing, pp. 55–60.

Dubois, D., Kerre, E., Mesiar, R., Prade, H. (2000b). Fuzzy Interval Analysis. In: Dubois, D., Prade, H. (Eds.), *Fundamentals of Fuzzy Sets*. Kluwer Academic Press, pp. 55–72.

Dubois, D., Moral, S., Prade, H. (1997). Semantics for possibility theory based on likelihoods. *Journal of Mathematical Analysis and Applications* **205**, 359–380.

Dwayer, P.S. (1951). *Linear Computations*. Wiley, New York.

Ely, J.S. (1993a). The VPI software package for variable precision interval arithmetic. *Interval Computation* **2** (2), 135–153.

Ely, J.S., Baker, G.R. (1993b). High-precision calculations of vortex sheet motion. *Journal of Computational Physics* **111**, 275–281.

Ferson, S. (2002). *RAMAS Risk Calc 4.0 Software: Risk Assessment with Uncertain Numbers*. Lewis Publishers, Boca Raton, FL.

Ferson, S., Kreinovich, L.R., Ginzburg, V., Sentz, K., Myers, D.S. (2003). Constructing probability boxes and Dempster-Shafer structures. Technical Report SAND2002-4015. Sandia National Laboratories, Albuquerque, NM.

Ferson, S., Kreinovich, V. (2005). Combining interval and probabilistic uncertainty: Foundations, algorithms, challenges—an overview. World-Wide Web.

Fischer, P.C. (1958). Automatic propagated and round-off error analysis. *Proceedings of the 13th national meeting of the Association for Computing Machinery, June 1958*. See http://portal.acm.org/citation.cfm?id=610971.

Fortin, J., Dubois, D., Fargier, H. (2006). Gradual numbers and their application to fuzzy interval analysis. *IEEE Transactions on Fuzzy Systems*, in press.

Frank, M.J., Nelsen, R.B., Schweizer, B. (1987). Best-possible bounds for the distribution of a sum—A problem of Kolmogorov. *Probability Theory and Related Fields* **74**, 199–211.

Gardeñes, E., Mielgo, H., Trepat, A. (1986). Modal intervals: Reasons and ground semantics. In: Nickel, K. (Ed.), *Interval Mathematics 1985*. Proceedings of the International Symposium, Freiburg, Federal Republic of Germany, September 23–26, 1985, Springer-Verlag, Berlin, pp. 27–35.

Guderley, K.G., Keller, C.L. (1972). A basic theorem in the computation of ellipsoidal error bounds. *Numerische Mathematik* **19** (3), 218–229.

Hales, T.C. (2000). Cannonballs and honeycombs. *Notices of the American Mathematical Society* **47**, 440–449.

Hansen, E.R. (1975). A generalized interval arithmetic. In: Nickel, K. (Ed.), *Interval Mathematics*, *Lecture Notes in Computer Science*, vol. **29**. Springer-Verlag, New York, pp. 7–18.

Hansen, E.R. (1978). Interval forms of Newton's method. *Computing* **20**, 153–163.

Hansen, E.R. (1980). Global optimisation using interval analysis—The multidimensional case. *Numerische Mathematik* **34**, 247–270.

Hansen, E.R. (1992). *Global Optimization Using Interval Analysis*. Marcel Dekker, New York.

Hansen, E.R. (2001). Publications related to early interval work of R.E. Moore, August 13, 2001. See http://interval.louisiana.edu/Moores_early_papers/bibliography.html.

Hanss, M. (2005). *Applied Fuzzy Arithmetic*. Springer-Verlag, Berlin.

Jahn, K. (1980). The importance of 3-valued notions for interval mathematics. In: Nickel, K.E. (Ed.), *Interval Mathematics*. Academic Press, New York, pp. 75–98.

Jamison, K.D., Lodwick, W.A. (2002). The construction of consistent possibility and necessity measures. *Fuzzy Sets and Systems* **132** (1), 1–10.

Jamison, K.D., Lodwick, W.A. (2004). Interval-valued probability measures. UCD/CCM Report No. 213.

Jamison, K.D., Lodwick, W.A. (2006) Interval-valued probability in the analysis of problems that contain a mixture of fuzzy, possibilistic and interval uncertainty. *International Journal of Approximate Reasoning*, in press.

Jamison, K.D., Lodwick, W.A., Kawai, M. (2002). A simple closed form estimation for the cumulative distribution function of a monotone function of random variables. UCD/CCM Report No. 187.

Jaulin, L. (2001a). Path planning using intervals and graphs. *Reliable Computing* **7** (1), 1–15.

Jaulin, L., Kieffer, M., Didrit, O., Walter, E. (2001b). *Applied Interval Analysis*. Springer-Verlag, New York.

Kahan, W.M. (1968a). Circumscribing an ellipsoid about the intersection of two ellipsoids. *Canadian Mathematical Bulletin* **11** (3), 437–441.

Kahan, W.M. (1968b). A more complete interval arithmetic. Lecture Notes for a Summer Course at University of Michigan.

Kaplan, S. (1981). On the method of discrete probability distributions in risk and reliability calculations—Applications to seismic risk assessment. *Journal of Risk* **1** (3), 189–196.

Kaucher, E. (1973). Über metrische und algebraische Eigenschaften eiginger beim numerischen Rechnen auftretender Räume. PhD thesis, University of Karlsruhe.

Kaucher, E. (1980). Interval analysis in the extended space $I\mathbb{R}$. *Computing (Suppl.)* **2**, 33–49.

Kaucher, E., Rump, S.M. (1982). E-methods for fixed point equations $f(x) = x$. *Computing* **28**, 31–42.

Kaufmann, A., Gupta, M.M. (1985). *Introduction to Fuzzy Arithmetic: Theory and Applications*. Van Nostrand Reinhold, New York.

Kawai, M. (2006). Discrete clouds. Seminar presentation April 28, 2006, Department of Mathematical Sciences of University of Colorado at Denver.

Kaymak, U., Sousa, J.M. (2003). Weighting of constraints in fuzzy optimization. *Constraints* **8**, 61–78.

Kearfott, R.B. (1996a). *Rigorous Global Search: Continuous Problem*. Kluwer Academic Publishers, Boston, MA.

Kearfott, R.B. (1996b). Interval computations: Introduction, uses, and resources. *Euromath Bulletin* **2** (1), 95–112.

Klaua, D. (1969). Partielle Mengen und Zhlen. *Mtber. Dt. Akad. Wiss.* **11**, 585–599.

Klir, G.J., Yuan, B. (1995). *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice Hall, Upper Saddle River, NJ.

Klir, G.J. (1997). Fuzzy arithmetic with requisite constraints. *Fuzzy Sets and Systems* **91** (2), 165–175.

Kolmogorov, A.N. (1941). Confidence limits for an unknown distribution function. *Annuals of Mathematics Statistics* **12**, 461–463.

Korenerup, P., Matula, D.W. (1983). Finite precision rational arithmetic: An arithmetic unit. *IEEE Transactions on Computers* **32** (4), 378–388.

Kulisch, U. (1971). An axiomatic approach to rounded computations. *Numerische Mathematik* **18**, 1–17.

Kulisch, U., Miranker, W.L. (1981). *Computer Arithmetic in Theory and Practice*. Academic Press, New York.

Kulisch, U., Miranker, W.L. (1986). The arithmetic of the digital computer: A new approach. *SIAM Review* **28** (1), 1–40.

Lin, T.Y. (2005). A function theoretic view of fuzzy sets: New extension principle. In: Filev, D., Ying, H. (Eds.), *Proceedings of NAFIPS 05*.

Lodwick, W.A. (1989). Constraint propagation, relational arithmetic in AI systems and mathematical programs. *Annals of Operations Research* **21**, 143–148.

Lodwick, W.A. (1999). Constrained interval arithmetic. *CCM Report* **138**.

Lodwick, W.A. (2002). Special Issue on theLinkages Between Interval Analysis and Fuzzy Set Theory. *Reliable Computing* **8**. (entire issue).

Lodwick, W.A., Jamison, K.D. (2003a). Special Issue on the Interfaces Between Fuzzy Set Theory and Interval Analysis. *Fuzzy Sets and Systems* **135**. (entire issue).

Lodwick, W.A., Jamison, K.D. (2003b). Estimating and validating the cumulative distribution of a function of random variable: Toward the development of distribution arithmetic. *Reliable Computing* **9**, 127–141.

Lodwick, W.A., Jamison, K.D. (2006). Interval-valued probability in the analysis of problems that contain a mixture of fuzzy, possibilistic and interval uncertainty. In: Demirli, K., Akgunduz, A. (Eds.), *2006 Conference of the North American Fuzzy Information Processing Society*, June 3–6, 2006, Montréal, Canada.

Makino, K., Berz, M. (2003). Taylor models and other validated functional inclusion methods. *International Journal of Pure and Applied Mathematics* **4** (4), 379–456.

Makino, K., Berz, M. (2005). Verified global optimization with Taylor model based range bounders. *Transactions on Computers* **11** (4), 1611–1618.

Mayer, G. (1995). Epsilon-inflation in verification algorithms. *Journal of Computational and Applied Mathematics* **60**, 147–169.

Mayer, G. (1996). Success in epsilon-inflation. In: Alefeld, G., Lang, B. (Eds.), *Scientific Computing and Validated Numerics*. Proceedings of the International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics SCAN-95, Wuppertal, Germany, September 26–29, 1995, Akademie-Verlag, Berlin, MA, pp. 98–104.

Moore, R.E. (1959a). Automatic error analysis in digital computation. Technical Report LMSD-48421. Lockheed Missile and Space Division. Sunnyvale, CA. See http://interval.louisiana.edu/Moores_early_papers/bibliography.html.

Moore, R.E., Yang, C.T. (1959b). Interval analysis I. Technical Report Space Div. Report LMSD285875. Lockheed Missiles and Space Co., Sunnyvale, CA.

Moore, R.E., Strother, W., Yang, C.T. (1960). Interval integrals. Technical Report Space Div. Report LMSD703073. Lockheed Missiles and Space Co., Sunnyvale, CA.

Moore, R.E. (1962). Interval arithmetic and automatic error analysis in digital computing. PhD thesis, Stanford University, Stanford, California, Published as Applied Mathematics and Statistics Laboratories Technical Report No. 25, November 15, 1962. See http://interval.louisiana.edu/Moores_early_papers/bibliography.html.

Moore, R.E. (1965). The automatic analysis and control of error in digital computing based on the use of interval numbers. In: Rall, L.B. (Ed.), *Error in Digital Computation, vol. I*. John Wiley & Sons, New York, pp. 61–130.
Moore, R.E. (1966). *Interval Analysis*. Prentice-Hall, Englewood Cliffs, NJ.
Moore, R.E. (1979). *Methods and Applications of Interval Analysis*. SIAM, Philadelphia, PA.
Moore, R.E. (1984). Risk analysis without Monte Carlo methods. *Freiburger Intervall-Berichte* **1**, 1–48.
Moore, R.E. (1992). Computing to arbitrary accuracy. In: Bresinski, C., Kulisch, U. (Eds.), *Computational and Applied Mathematics I: Algorithms and Theory*. North-Holland, Amsterdam, pp. 327–336.
Moore, R.E. (1999). The dawning. *Reliable Computing* **54**, 423–424.
Moore, R.E., Lodwick, W.A. (2003). Interval analysis and fuzzy set theory. *Fuzzy Sets and Systems* **135** (1), 5–9.
Nahmias, S. (1978). Fuzzy variable. *Fuzzy Sets and Systems* **1**, 97–110.
Nelsen, R.B. (1995). Copulas, characterization, correlation, and counterexamples. *Mathematics Magazine* **68** (3), 193–198.
Negoita, C.V., Ralescu, D.A. (1975). *Applications of Fuzzy Sets to Systems Analysis*. Birkhäuser, Boston, MA.
Neumaier, A. (1990). *Interval Methods for Systems of Equations*. Cambridge University Press, Cambridge.
Neumaier, A. (1993). The wrapping effect, ellipsoid arithmetic, stability and confidence regions. *Computing Supplementum* **9**, 175–190.
Neumaier, A. (2004a). Clouds, fuzzy sets and probability intervals. *Reliable Computing* **10**, 249–272.
Neumaier, A. (2004b). Complete search in continuous global optimization and constraint satisfaction. In: Iserles, A. (Ed.), *Acta Numerica 2004*. Cambridge University Press, Cambridge, pp. 271–369.
Neumaier, A. (2005). Structure of clouds (submitted for publication). See http://www.mat.univie.ac.at/~neum/papers.html.
Nguyen, H.T. (1978). A note on the extension principle for fuzzy sets. *Journal of Mathematical Analysis and Applications* **64**, 369–380.
Nickel, K. (1969). Triplex-Algol and its applications. In: Hansen, E.R. (Ed.), *Topics in Interval Analysis*. Oxford University Press, pp. 10–24.
Olsen, G. (2005). The inverse probability method: An interval algorithm for measuring uncertainty. Masters Thesis, University of Colorado at Denver, Department of Mathematics.
Ortolf, H.-J. (1969). *Eine Verallgemeinerung der Intervallarithmetik*. Gesellschaft für Mathematik und Datenverarbeitung, Bonn. Nr. 11, pp. 1–71.
Phillips, G.M. (1981). Archimedes the numerical analyst. *American Mathematical Monthly* **81**, 165–169.

Popova, E.D. (1998). See http://www.math.bas.bg/~epopova/directed.html.
Pryce, J.D., Corliss, G.F. (2006). Interval arithmetic with containment sets. (Submitted for publication.).
Ramik, J. (1986). Extension principle in fuzzy optimization. *Fuzzy Sets and Systems* **19**, 29–35.
Ratschek, H., Rokne, J. (1988). *New Computer Methods for Global Optimization*. Ellis Horwood, Chichester, England.
Revol, N., Rouillier, F. (2005). Motivations for an arbitrary precision interval arithmetic and the MPFI library. *Reliable Computing* **11** (4), 275–290.
Sáinz, M.A. (2001). Modal intervals. *Reliable Computing* **7** (2), 77–111.
Schulte, M.J., Swartzlander, E.E. Jr. (2000). A family of variable-precision interval processors. *IEEE Transactions on Computers* **49** (5), 387–397.
Shafer, G. (1987). Belief functions and possibility measures. In: Bezdek, J.C. (Ed.), *Analysis of Fuzzy Information: Mathematics and Logic, vol. 1*. CRC Press, Boca Raton, FL, pp. 51–84.
Smale, S. (1998). Mathematical problems for the next century. *Mathematical Intelligencer* **20** (2), 7–15.
Stolfi, J., Andrade, M.V.A., Comba, J.L.D. (1994). Affine arithmetic: A correlation-sensitive variant of interval arithmetic. See http://www.dcc.unicamp.br/~stolfi/EXPORT/projects/affine-arith.
Springer, M.D. (1979). *The Algebra of Random Variables*. Wiley & Sons, New York.
Stetter, H.J. (1978). The defect correction principle and discretization methods. *Numerische Mathematik* **29**, 425–443.
Strother, W. (1952). Continuity for multi-valued functions and some applications to topology. Doctoral dissertation, Tulane University.
Strother, W. (1955). Fixed points, fixed sets, and $m$-retracts. *Duke Mathematical Journal* **22** (4), 551–556.
Strother, W. (1958). Continuous multi-valued functions. *Boletim da Sociedade de Matematica de São Paulo* **10**, 87–120.
Sunaga, T. (1958). Theory of an interval algebra and its application to numerical analysis. *RAAG Memoirs* **2**, 547–564. Downloaded http://www.cs.utep.edu/interval-comp/early.html.
Tonon, F. (2004). On the use of random set theory to bracket the results of Monte Carlo simulations. *Reliable Computing* **10**, 107–137.
Tucker, W. (2002). A rigorous ODE solver and Smale's 14th problem. *Foundation of Computational Mathematics* **2**, 53–117.
Tupper, J.A. (1996). Graphing equations with generalized interval arithmetic. PhD thesis, University of Toronto.
Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. Chapman & Hall, London.

Walley, P. (1999). Towards a unified theory of imprecise probability. In: De Cooman, G., Cozman, F.G., Moral, S., Walley, P. (Eds.), *Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications*, held at the Conference Center "Het Pand" of the Universiteit Gent, Ghent, Belgium, 29 June–2 July 1999, ISIPTA.

Walster, G.W. (1998). The extended real interval system. (Personal copy from the author).

Wang, Z., Klir, G.J. (1992). *Fuzzy Measure Theory*. Plenum Press, New York.

Warmus, M. (1956). Calculus of approximations. *Bulletin de l'Académie Polonaise de Sciences* **3** (4), 253–259. Downloaded from http://www.cs.utep.edu/interval-comp/early.html.

Warmus, M. (1961). Approximations and inequalities in the calculus of approximations. Classification of approximate numbers. *Bulletin de l'Académie Polonaise de Sciences* **4** (4), 241–245. Downloaded from http://www.cs.utep.edu/interval-comp/early.html.

Weichselberger, K. (1996). Interval probability on finite sample spaces. In: Rieder, H. (Ed.), *Robust Statistics, Data Analysis and Computer-Intensive Methods*. Springer-Verlag, New York, pp. 391–409.

Weichselberger, K. (2000). The theory of interval-probability as a unifying concept for uncertainty. *International Journal of Approximate Reasoning* **24**, 149–170.

Williamson, R.C., Downs, T. (1990a). Probabilistic arithmetic I: Numerical methods for calculating convolutions and dependency bounds. *International Journal of Approximate Reasoning* **4**, 89–158.

Williams, R.C. (1990b). Interval arithmetic and probability arithmetic. In: Ullrich, C. (Ed.), *Computer Arithmetic and Self-Validating Numerical Methods*. Academic Press, Boston, MA, pp. 67–80.

Yager, R.R. (1986). A characterization of the extension principle. *Fuzzy Sets and Systems* **18**, 205–217.

Young, R.C. (1931). The algebra of many-valued quantities. *Mathematische Annalen* **104**, 260–290.

Zadeh, L.A. (1965). Fuzzy sets. *Information and Control* **8**, 338–353.

Zadeh, L.A. (1968). Probability measures of fuzzy events. *Journal of Mathematical Analysis and Applications* **23**, 421–427.

Zadeh, L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning. *Information Sciences*. Part I: **8**, 199–249, Part II: **8**, 301–357, Part III: **9**, 43–80.

Zadeh, L.A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems* **1**, 3–28.

Zadeh, L.A. (2005). Toward a generalized theory of uncertainty (GTU)—An outline. *Information Science* **172** (1–2), 1–40.