Proceedings of ASME 2010 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2010
August 15 – 18, 2010, Montreal, Quebec, Canada

DETC2010/28790

MULTISCALE VARIABILITY AND UNCERTAINTY QUANTIFICATION BASED ON A GENERALIZED MULTISCALE MARKOV MODEL

Yan Wang
Woodruff School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332

ABSTRACT

Variability is inherent randomness in systems, whereas uncertainty is due to lack of knowledge. In this paper, a generalized multiscale Markov (GMM) model is proposed to quantify variability and uncertainty simultaneously in multiscale system analysis. The GMM model is based on a new imprecise probability theory that has the form of generalized interval, which is a Kaucher or modal extension of classical set-based intervals to represent uncertainties. The properties of the new definitions of independence and Bayesian inference are studied. Based on a new Bayes' rule with generalized intervals, three cross-scale validation approaches that incorporate variability and uncertainty propagation between scales are also developed.

1. INTRODUCTION

Multiscale systems are systems consisting of hierarchical structures with different sizes and exhibit patterns of behaviors as the diagnostics of interactions among subsystems at lower levels recursively. Human cells, atmospheric turbulence, ecosystems, product-materials hierarchies, global supply chain, etc. are such examples. In modeling and analyzing multiscale systems, the effects of variability and uncertainty should be studied. Variability and uncertainty are artifacts that exist universally. Variability is the inherent randomness due to fluctuation and perturbation. In literature, variability is also referred to as stochastic uncertainty, simulation uncertainty, aleatory uncertainty, and irreducible uncertainty. This component is irreducible even by additional measurements. Uncertainty is due to lack of perfect knowledge or enough information about the system. Uncertainty is also known as epistemic uncertainty, reducible uncertainty, and model form uncertainty. Since uncertainty is caused by the lack of information about a system, it can be reduced by increasing our knowledge to fill the information gap.

There are many arguments (e.g., [1,2,3]) that support the separation between uncertainty and variability. Uncertainties have different sources, including lack of data or missing data; conflicting information if there are multiple sources of information causing inconsistency; conflicting beliefs of domain experts when data are not available and the analyst has to build models based on experts' judgements and opinions; lack of introspection of systems under studied; measurement errors and numerical errors; and lack of evidence about the dependency among factors and variables; etc.

Most of the existing stochastic models only focus on one length scale. For multiscale systems, variability and uncertainty usually propagate between scales and are inter-dependent. For instance, distributions of defects in alloy crystals determine the reliability of structures. Physical properties of materials are manifestations of atomic-level electron densities and distributions. We need multiscale stochastic models to incorporate uncertainties and capture cross-scale correlation.

Probability theory provides a common ground to quantify both variability and uncertainty and so far is the most popular approach. As a result, in some literature, epistemic probability and aleatory probability are named. Traditionally probabilistic properties are quantified by precise values of probability measures and their parameters (e.g. means and higher-order moments). However, the precise probability theory has limitations to represent uncertainty. The most significant one is that it does not differentiate the *total ignorance* from other probability distributions. The total ignorance means that there is absolutely no information about the system or subject under study. Based on the principle of maximum entropy, uniform distributions are usually assumed when the precise probability theory is applied in this case. The problem arises because introducing any form of distribution itself has introduced extra knowledge in. "Knowing unknown" is not the total ignorance. Another limitation of precise probability to quantify uncertainty is to represent *indeterminacy* and *inconsistency*. When no data is available and people have limited ability to determine their own subjective probabilities, or subjective probabilities from

different people are inconsistent, precise probabilities do not capture a range of opinions or estimations adequately without assuming some consensus of precise values on the distribution of opinions. "Agreeing disagreed" is not the best way to capture inconsistency.

Here, we propose to use imprecise probabilities to represent uncertainty and variability. Instead of a precise value of the probability $P(E) = p$ associated with an event E, a pair of lower and upper probabilities $P(E) = [\underline{p}, \overline{p}]$ are used to include a *set* of probabilities and quantify the uncertainty. Imprecise probability differentiates uncertainty from variability both qualitatively and quantitatively, which is the alternative to the traditional sensitivity analysis in probabilistic reasoning to model imprecision. The range of the interval $[\underline{p}, \overline{p}]$ captures the uncertainty component and indeterminacy. $P = [0,1]$ accurately represents the total ignorance. When $\underline{p} = \overline{p}$, the degenerated interval probability becomes a precise one without uncertainty. In a general sense, imprecise probability is a *generalization* of precise probability.

We say a modeling or simulation mechanism is *reliable* if the estimation from analysis or simulation is both *complete* and *sound* with respect to uncertainties. A *complete* range estimation of possible values includes all possible occurrences. A *sound* range estimation does not include impossible occurrences. The purpose of using imprecise probability in system analysis is to improve the robustness of prediction and support more informed decision. The *robustness* can be generally measured by the chance or probability that a range estimation of uncertainty includes all possible occurrences. An efficient and verifiable way to establish the confidence of completeness and soundness in modeling and simulation (M&S) is important. To make informed and robust decisions, we usually need to know the worst-case and best-case scenarios among a range of other possible outcomes. The existing sampling-based simulation mechanism such as second-order Monte Carlo and sensitivity analysis cannot provide such information efficiently. The existing uncertainty quantification methods are not general enough to model the total uncertainty. There is a need of generic mathematical framework to study the total uncertainty in multiscale complex systems. Therefore, in this paper we propose a generalized multiscale Markov (GMM) model to model the total uncertainty and support reliable simulation.

In the remainder of the paper, we first give a brief overview of relevant work in multiscale simulation, variability and uncertainty quantification, imprecise probability, and generalized interval in Section 2. In Section 3, the new imprecise probability theory based on the generalized intervals is described. In Section 4, the generalized multiscale Markov Model is proposed. Three new cross-scale validation approaches are developed and demonstrated in Section 5.

2. BACKGROUND

2.1 Stochastic Models to Simulate with Variabilities

Various stochastic models to accommodate variabilities have been developed in different scales. At the traditional macro- or bulk-scale of engineering, stochastic or probabilistic finite element analysis with random fields has been extensively studied. The basic idea is to incorporate variabilities of geometries, material properties, and loads in finite element analysis. The research issues include how to estimate variance of performance (e.g., displacement, strain, and stress) given the variabilities of inputs [4]; how to perform variance estimations with linear and higher-order approximations [5]; how to improve the computational efficiency in spectral approximations by the Karhunen-Loève (K-L) decomposition [6] and its generalizations and extensions [7,8,9]; how to find optimal solutions under the constraint of reliability [10]; and others.

At the meso-scale, dislocation dynamics is a popular tool to simulate plastic deformation of crystalline structures, where Newtonian-like equations are used to describe the motion of crystal defects within the stress fields. Extended from deterministic models, stochasticity was recently introduced in dislocation dynamics simulations to incorporate the fluctuation effects of internal stress [11] and spatial distributions [12,13] caused by long-range dislocation interaction, and thermal dissipation [14] during plastic flow. The stochastic DD problem then is formulated and solved as Langevin-type evolution equations.

The models reviewed above only consider stochasticity within one scale. Assumptions are made such that variabilities among different lengths scales are separable. That is, the randomness at macro-scale is independent from that of micro-scale. This "homogenization" approach does not always model the real world. For example, the effective variance of moduli obtained by averaging over small domains of composite materials does not agree with the one obtained by a sufficiently large representative volume element. Furthermore, damage and fracture are highly sensitive to very local defects [15]. Decoupling variational information between length scales will compromise the accuracy of predictions.

2.2 Multiscale Simulation under Uncertainty

Plenty of research has been done on deterministic multiscale simulation. Relatively, few research is focused on stochastic information integration. Recently, Choi et al. [16] represented variabilities as multiscale Gaussian models on a pyramid graph structure. Multiscale information assimilation was achieved by a so-called walk-sum analysis for both long-range and local dependencies. As an extension of Arlequin coupling framework, Chamoin et al. [17] proposed a stochastic coupling approach based on homogenization of material properties between length scales for Monte Carlo simulation. Ganapathysubramanian and Zabaras [18] developed an upscaling approach to derive coarse-scale probability distributions from fine-scale distributions based on sampling in low-dimensional space. Arnst and Ghanem [19] took another upscaling approach to approximate fine-scale probability

distributions by the K-L decomposition. Chen and co-workers [20,21] also developed an upscaling approach based on the K-L decomposition and integrated it with stochastic finite-element analysis.

The above methods are intended to solve the issue of multiscale variability information exchange. Domain specific assumptions of probability distributions were made so that analysis is computationally tractable. More importantly, uncertainty and variability were not differentiated in these methods. Consequently, the effects of lack of information verses fluctuation are indistinguishable. Given the very different nature of variability and uncertainty, independent quantification of the two will be very useful to understand the analysis results and make appropriate decisions accordingly.

In summary, most of the existing stochastic models except for a few only focus on one length scale. There are no multiscale stochastic models that consider uncertainties explicitly. Completeness and soundness need to be verified rigorously in uncertainty quantification. The proposed GMM model allows us to quantify cross-scale dependency and information loss during exchanges between scales.

2.3 Hidden Markov Models

The proposed GMM model is a *generalization* of hidden Markov models (HMM) to consider hierarchical complex systems and uncertainties related to models and parameters. The HMM [22,23] is an extension of regular Markov chain, where the state variables are not directly observable. Instead, all statistical inference about the Markov chain itself has to be done in terms of observable variables. HMM has been applied in many fields that are based on the analysis of discrete-valued time series, such as speech recognition, genetic profile and classification. Notice that a special case of hidden Markov models is the *Gaussian linear state-space* model, where state series are linearly dependent on history, subject to process white noises. And observations are also linearly dependent on states, subject to measurement noises. The linear state-space model is also known as the *Kalman filter*. It is has been widely used in many fields in science and engineering. However, it considers only linear relationships. HMM is more general.

There has been some research to extend HMM to hierarchical systems. For instance, the multiresolution hidden Markov model was used to recursively represent natural language utterance, in which each state is a hidden Markov model itself that has a sequence of state transitions [24]. Tree-structured hierarchical Markov models were employed in pattern recognition and classification of images [25,26]. Various improvement and extensions have been proposed since then. Again, uncertainty information is not explicitly captured in these Markov models. From the perspective of uncertainty quantification, imprecise probability we propose is a generalization of traditional precise probability.

2.4 Imprecise Probability

Our proposed approach uses imprecise probabilities to quantify variability and uncertainty simultaneously. Many representations of imprecise probabilities have been developed. For example, the Dempster-Shafer evidence theory [27,28] characterizes evidence with discrete probability masses associated with a power set of values, where Belief-Plausibility pairs are used to measure uncertainties. The behavioral imprecise probability theory [1] models uncertainties with the *lower prevision* (supremum acceptable buying price) and the *upper prevision* (infimum acceptable selling price) with behavioral interpretations. The possibility theory [29] represents uncertainties with Necessity-Possibility pairs. A random set [30] is a multi-valued mapping from the probability space to the value space. Probability bound analysis [31] captures uncertain information with pairs of lower and upper distribution functions. F-probability [32] incorporates intervals into probability values which maintain the Kolmogorov properties. Fuzzy probability [33] considers probability distributions with fuzzy parameters. A cloud [34] is a combination of fuzzy sets, intervals, and probability distributions.

One common problem of the above set-based imprecise probability theories is that the calculation is cumbersome. Linear and nonlinear optimization methods are dependent upon to search lower and upper bounds during reasoning. Different from them, we recently proposed an imprecise probability with a generalized interval form [35,36], where the probabilistic calculus structure is greatly simplified based on the algebraic properties of the Kaucher arithmetic [37] for generalized intervals.

2.5 Interval Arithmetic and Generalized Interval

In the interval arithmetic [38], it is guaranteed that the output intervals calculated from the arithmetic include all possible combinations of real values within the respective input intervals. That is, if $[\underline{x},\overline{x}]$ and $[\underline{y},\overline{y}]$ are two real intervals (i.e., $\underline{x},\overline{x},\underline{y},\overline{y}\in\mathbb{R}$) and let $\circ\in\{+,-,\times,/\}$, then we have $\forall x\in[\underline{x},\overline{x}],\forall y\in[\underline{y},\overline{y}],\exists z\in[\underline{x},\overline{x}]\circ[\underline{y},\overline{y}],x\circ y=z$. For example, $[1,3]+[2,4]=[3,7]$ guarantees that $\forall x\in[1,3],\forall y\in[2,4],\exists z\in[3,7],x+y=z$. Similarly, $[3,7]-[1,3]=[0,6]$ guarantees that $\forall x\in[3,7],\forall y\in[1,3],\exists z\in[0,6],x-y=z$. This is an important property that ensures the completeness of range estimations. When input variables are not independent, the output results will over-estimate the actual ranges. This only affects the soundness of estimations, not completeness. Some special techniques also have been developed to avoid over-estimations based on monotonicity properties of functions.

Generalized interval [39,40] is an extension of the above set-based classical interval with better algebraic and semantic properties based on the Kaucher arithmetic [37]. A generalized interval $\mathbf{x}:=[\underline{x},\overline{x}]\ (\underline{x},\overline{x}\in\mathbb{R})$ is not constrained by $\underline{x}\le\overline{x}$ any more. Therefore, $[4,2]$ is also a valid interval and called *improper*, while the traditional interval is called *proper*. Based on the Theorems of Interpretability [39], generalized interval provides more semantic power to help verify completeness and

soundness of range estimations by logic interpretations.

The four examples in Table 1 illustrate the interpretations for operator "+", where the range estimation $[\underline{z},\overline{z}]=[4,7]$ in the 1st row is *complete* and the estimation $[\underline{z},\overline{z}]=[7,4]$ in the 4th row is *sound*. $-,\times,/$ have the similar semantic properties. More information of generalized interval can be found in [41,42,43].

In the proposed model, uncertainty propagation will be based on both the interval arithmetic and the Kaucher arithmetic. This allows us to interpret interval results so that the completeness and soundness can be verified rigorously.

Compared to the *semi-group* formed by the classical set-based intervals, generalized intervals form a *group*. Therefore, arithmetic operations of generalized intervals are simpler. The set of generalized intervals is denoted by $\mathbb{KR}=\{[\underline{x},\overline{x}]\mid \underline{x},\overline{x}\in\mathbb{R}\}$. The set of proper intervals is $\mathbb{IR}=\{[\underline{x},\overline{x}]\mid \underline{x}\le\overline{x}\}$, and the set of improper interval is $\overline{\mathbb{IR}}=\{[\underline{x},\overline{x}]\mid \underline{x}\ge\overline{x}\}$. The relationship between proper and improper intervals is established with the operator *dual* as

$$\mathrm{dual}[\underline{x},\overline{x}]:=[\overline{x},\underline{x}] \tag{2.1}$$

The *less than or equal to* partial order relationship between two generalized intervals is defined as

$$[\underline{x},\overline{x}]\le[\underline{y},\overline{y}]\Leftrightarrow \underline{x}\le\underline{y}\wedge\overline{x}\le\overline{y} \tag{2.2}$$

With the Kaucher arithmetic, generalized intervals form a lattice structure similar to real arithmetic, which is not available in the classical interval arithmetic. This property significantly simplifies the computational requirement. For instance, in classical interval arithmetic, $[0.2,0.3]+[0.2,0.4]=[0.4,0.7]$. However, $[0.4,0.7]-[0.2,0.3]=[0.1,0.5]\neq[0.2,0.4]$. Furthermore, $[0.1,0.2]-[0.1,0.2]=[-0.1,0.1]\neq 0$. In the Kaucher arithmetic, if a *dual* is associated with "–", then $[0.4,0.7]-\mathrm{dual}[0.2,0.3]=[0.4,0.7]-[0.3,0.2]=[0.2,0.4]$. $[0.1,0.2]-\mathrm{dual}[0.1,0.2]=0$. "×" and "÷" are similar.

3. IMPRECISE PROBABILITY WITH THE GENERALIZED INTERVAL FORM

3.1 Basics

Definition 1. Given a sample space Ω and a σ-algebra $\mathcal{A}$ of random events over Ω, the generalized interval probability $\mathbf{p}\in\mathbb{KR}$ is defined as $\mathbf{p}:\mathcal{A}\to[0,1]\times[0,1]$ which obeys the axioms of Kolmogorov: (1) $\mathbf{p}(\Omega)=[1,1]$; (2) $[0,0]\le\mathbf{p}(E)\le[1,1]\ (\forall E\in\mathcal{A})$; and (3) for any countable mutually disjoint events $E_i\cap E_j=\varnothing\ (i\neq j)$, $\mathbf{p}\left(\bigcup_{i=1}^{n}E_i\right)=\sum_{i=1}^{n}\mathbf{p}(E_i)$. Here "$\le$" is defined as in Eq.(2.2).

Definition 2. The probability of *union* is defined as $\mathbf{p}(A):=\sum_{S\subseteq A}(-\mathrm{dual})^{|A|-|S|}\mathbf{p}(S)$ for $A\subseteq\Omega$.

The most important property of the generalized interval probability is the *logic coherence constraint* (LCC): For a mutually disjoint event partition $\bigcup_{i=1}^{n}E_i=\Omega$, $\sum_{i=1}^{n}\mathbf{p}(E_i)=1$. The LCC ensures that the imprecise probabilities are logically coherent with precise probabilities. For instance, given that $\mathbf{p}(down)=[0.2,0.3]$, $\mathbf{p}(idle)=[0.3,0.5]$, $\mathbf{p}(busy)=[0.5,0.2]$ for a system's working status, we can interpret it as

$$(\forall p_1\in[0.2,0.3])(\forall p_2\in[0.3,0.5])(\exists p_3\in[0.2,0.5])\ (p_1+p_2+p_3=1)$$

Accordingly, we differentiate *non-focal* events ("*busy*" in this example) from *focal* events ("*down*", "*idle*"). An event E is focal if the associated semantics for $\mathbf{p}(E)$ is universal. Otherwise, it is a non-focal if the semantics is existential. While the uncertainties associated with focal events are critical to the analyst, those associated non-focal events are not.

Table 1: Illustrations of the semantic extension of generalized interval

Algebraic Relation: $[\underline{x},\overline{x}]+[\underline{y},\overline{y}]=[\underline{z},\overline{z}]$	**Corresponding Logic Interpretation**	**Quantifier of** $[\underline{z},\overline{z}]$	**Range Estimation of** $[\underline{z},\overline{z}]$
$[2,3]+[2,4]=[4,7]$	$(\forall x\in[2,3])(\forall y\in[2,4])(\exists z\in[4,7])(x+y=z)$	$\exists$	[4,7] is *complete*
$[2,3]+$**[4,2]**$=$**[6,5]**	$(\forall x\in[2,3])(\forall z\in[5,6])(\exists y\in[2,4])(x+y=z)$	$\forall$	[5,6] is *sound*
[3,2]$+[2,4]=[5,6]$	$(\forall y\in[2,4])(\exists x\in[2,3])(\exists z\in[5,6])(x+y=z)$	$\exists$	[5,6] is *complete*
[3,2]$+$**[4,2]**$=$**[7,4]**	$(\forall z\in[4,7])(\exists x\in[2,3])(\exists y\in[2,4])(x+y=z)$	$\forall$	[4,7] is *sound*

3.2 Conditional Probability and Conditional Independence

The concepts of conditional probability and independence are essential for the classical probability theory. With them, we can decompose a complex problem into simpler and manageable components. Similarly, they are critical for imprecise probabilities. However, there is no agreement on how to define them yet.

Different from all other forms of imprecise probabilities, which are based on convex probability sets, our conditional probability is defined directly from the marginal ones.

Definition 3. The *conditional probability* $\mathbf{p}(E \mid C)$ for all $E, C \in \mathcal{A}$ is defined as

$$\mathbf{p}(E \mid C) := \frac{\mathbf{p}(E \cap C)}{\operatorname{dual}\mathbf{p}(C)} = \left[\frac{\underline{p}(E \cap C)}{\underline{p}(C)}, \frac{\overline{p}(E \cap C)}{\overline{p}(C)} \right] \quad (2.3)$$

when $\mathbf{p}(C) > 0$.

Thanks to the nice algebraic properties of generalized intervals, this definition can greatly simplify computation in applications. In traditional imprecise probabilities, linear and nonlinear programming procedures are heavily dependent upon to compute convex hulls of probability sets. In our definition, only algebraic computation is necessary.

Definition 4. For $A, B, C \in \mathcal{A}$, A is said to be *conditionally independent* with B on C if and only if

$$\mathbf{p}(A \cap B \mid C) = \mathbf{p}(A \mid C)\mathbf{p}(B \mid C) \quad (2.4)$$

Definition 5. For $A, B \in \mathcal{A}$, A is said to be *independent* with B if and only if

$$\mathbf{p}(A \cap B) = \mathbf{p}(A)\mathbf{p}(B) \quad (2.5)$$

The independence in Definition 5 is a special case of conditional independence in Definition 4, where C is the complete sample space Ω. The conditional independence in Eq.(2.4) can also have a second form, as shown in Theorem 3.1.

Theorem 3.1. $\mathbf{p}(A \cap B \mid C) = \mathbf{p}(A \mid C)\mathbf{p}(B \mid C) \Leftrightarrow \mathbf{p}(A \mid B \cap C) = \mathbf{p}(A \mid C)$.

In addition to computational simplification, our approach also allows for logic interpretation. Eq.(2.4) is interpreted as

$$\left(\forall p_1 \in \mathbf{p}'(A \mid C)\right)\left(\forall p_2 \in \mathbf{p}'(B \mid C)\right)\left(\exists p_3 \in \mathbf{p}'(A \cap B \mid C)\right)\left(p_1 p_2 = p_3\right)$$

This is very useful to verify the completeness and soundness of interval bound estimations.

The most intuitive meaning of "independence" is that an independence relationship satisfies several *graphoid* properties. Three obvious grahoid properties of our new definition of independence are listed as follows, where "$\perp$" denotes independence.

Corollary 3.2 (Symmetry) For $X, Y, Z \in \mathcal{A}$, $X \perp Y \mid Z \Rightarrow Y \perp X \mid Z$.

Corollary 3.3 (Redundancy). For $X, Y \in \mathcal{A}$, $X \perp Y \mid X$.

Corollary 3.4 (Contraction). For $X, Y, Z, W \in \mathcal{A}$, $\left(X \perp Y \mid Z\right) \& \left(X \perp W \mid \left(Y \cap Z\right)\right) \Rightarrow X \perp \left(W \cap Y\right) \mid Z$.

4. GENERALIZED MULTISCALE MARKOV MODEL

We propose a new and generic probabilistic model to account for multiscale variability and uncertainty information in hierarchical systems. The proposed generalized multiscale Markov (GMM) model essentially captures spatial and temporal dependency. As illustrated in Figure 1, the spatial domains in three length scales Ω_x, Ω_y, and Ω_z ($\Omega_x \subset \Omega_y \subset \Omega_z$) are subdivided into cells. The state of each cell is represented as a random variable, denoted as x_i, y_j, z_k respectively at three scales. The state value of a cell is dependent on those values of neighboring cells. Neighboring cells are connected in the graphic model, whereas non-neighboring ones are not.

The *spatial dependency* is represented as a conditional probability. For instance, $\mathbf{p}\left(x_i = a \mid x_{i,1} = b_1, x_{i,2} = b_2, \ldots, x_{i,l} = b_l\right)$ is the probability that the state variable x_i has value of a given that its l *neighboring* cells have the respective state values of $\left(b_1, \ldots b_l\right)$. In the example of Figure 1, x_i has $l = 4$ neighbors. Notice that neighbors do not necessarily mean that they are spatially close. If long-range correlations exist, two neighboring cells could spatially far apart but statistically correlated.

Between different scales, there are also dependency relationships. The *scale dependency* is also represented as a conditional probability. For instance, in Figure 1, the state of cell y_j at Scale Y is dependent on the state values of corresponding subdomain Ω_x, i.e. $\mathbf{p}\left(y_j \mid x_1, \ldots, x_i, \ldots, x_9\right)$.

In general, the true state values of cells may or may not be directly observable. Theoretically, all observed values in experiments contain the effects of uncertainties and variabilities. Therefore, the observed states are just another set of random variables that are dependent on true state values. Here, the *observation dependency* is included in the GMM model. Without loss of generality, we assume that each of the cells in different scales has its corresponding observable state.

For instance, in Figure 1, the true states of the cells on the left-hand side are hidden, and the corresponding observations are on the right-hand side. The probability of observing $X_i = b$ given that $x_i = a$ is $\mathbf{p}\left(X_i = b \mid x_i = a\right)$. Similarly, we have $\mathbf{p}\left(Y_j \mid y_j\right)$ and $\mathbf{p}\left(Z_k \mid z_k\right)$ at other length scales. If there are states that are not observable, the number of observation dependency relationships is reduced.

The most important and unique generalization of the proposed GMM model is that imprecise probabilities based on generalized intervals are used in the model. With imprecise probabilities, both variability and uncertainty can be explicitly incorporated and analyzed. With generalized intervals, inference and reasoning can be significantly simplified. Therefore, the proposed model improves computational efficiency while gaining more information from simulation or analysis results. Notice that the model illustrated in Figure 1 shows spatial-dependency only. To capture time-dependency, state transitions can also be achieved. That is, a GMM model with one-dimensional neighborhood relationships will represent the state transition history within one cell in Figure 1.

We call our GMM model "generalized" because of three levels of generalizations. First, our multiscale Markov model is a generalization of commonly used Markov chains and hidden Markov models. Second, our Markov model with imprecise probabilities is a generalization of traditional models with precise probabilities. Third, our new form of imprecise probability based on generalized intervals is also a generalization of imprecise probabilities.

With the incorporation of imprecise probabilities with the generalized interval form, a concise form of interval probabilities similar to the traditional precise probability can be achieved. The most important properties are *localities*. These include the locality of observation and the locality of scale.

4.1 Locality of Observation

Theorem 4.1 For two disjoint subdomains $\mathcal{A}_i$ and $\mathcal{A}_j$ at Scale X, if the hidden states $x_{i\in\mathcal{A}_i}$ and $x_{j\in\mathcal{A}_j}$ are independent and the corresponding observations are also independent, then

$$\mathbf{p}\left(X_{\mathcal{A}_i}, X_{\mathcal{A}_j} \mid x_{i\in\mathcal{A}_i}, x_{j\in\mathcal{A}_j}\right) = \mathbf{p}\left(X_{\mathcal{A}_i} \mid x_{i\in\mathcal{A}_i}\right)\mathbf{p}\left(X_{\mathcal{A}_j} \mid x_{j\in\mathcal{A}_j}\right)$$

Theorem 4.1 provides the algebraic convenience to decompose a complex system into smaller subsystems. Independent experimental measurements can be performed without losing the grand picture of variabilities and uncertainties of the whole system. For instance, in analyzing the reliability of automobile cooling system, if the radiator and thermostat are independent, then the uncertainties associated with the probabilities of failures for two subassemblies can be estimated by experiments separately without affecting the total uncertainty of the system. The uncertainty associated with the reliability of the complete cooling system can be easily calculated by the individual ones from subassemblies.

4.2 Locality of Scale

Theorem 4.2 If $\mathcal{A}_i$, $\mathcal{B}_j$, $\mathcal{C}_k$ are subdomains at Scales X, Y, Z respectively with $\mathcal{A}_i \subset \mathcal{B}_j \subset \mathcal{C}_k$,

$$\mathbf{p}\left(x_{\mathcal{A}_i} \mid y_{\mathcal{B}_j}, z_{\mathcal{C}_k}\right) = \mathbf{p}\left(x_{\mathcal{A}_i} \mid y_{\mathcal{B}_j}\right)$$

Theorem 4.2 allows us to simplify multiscale variability and uncertainty analysis. The propagation of uncertain information between scales is only limited to two that are closely related. For instance, the imprecision of crystal orientations only affects the anisotropic property estimates of one grain. However, those associated with polycrystalline solids of further larger scale will not be affected.

The simplicity of the above locality properties is due to the definition of conditional probability in Eq.(2.3) as well as the group properties of the generalized interval.

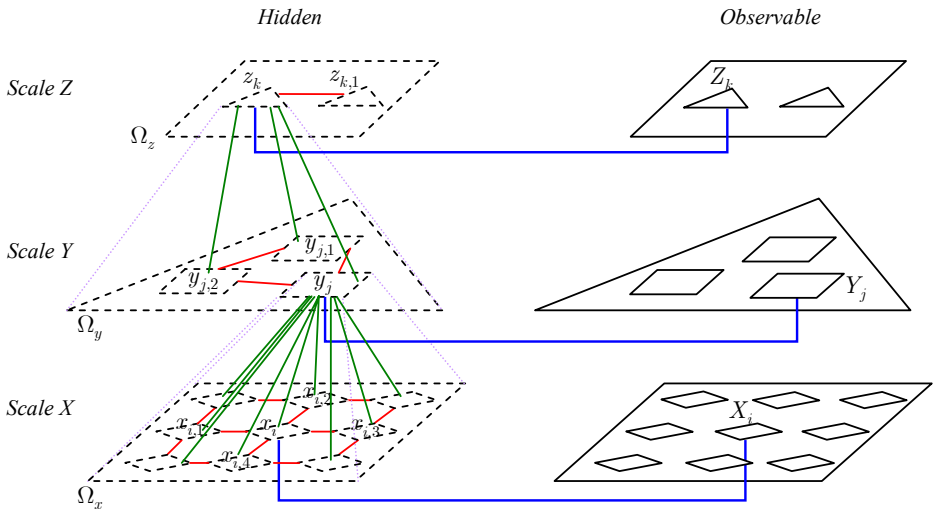

Figure 1: The proposed generalized multiscale Markov model for hierarchical systems to capture spatial and scale dependency

5. CROSS-SCALE VALIDATION

When small-scale (or large-scale) experiments are not possible, or the measurements are not feasible or reliable at one particular scale, we may conduct experiments at a larger (or smaller) scale to measure system properties to validate models or assumptions. For instance, in functional nano-materials design, instead of direct measuring molecule-level properties which is usually expensive or even impossible, the measurement of aggregated properties at macro-scale can be easier and much more accurate. In contrast, it is impossible to measure global temperature change. We only depend on regional ocean water temperature changes to predict the global picture. Cross-scale validation thus is an important tool in studying multiscale systems.

Definition 5. The *Bayes' rule with generalized intervals* (GIBR) is defined as

$$\mathbf{p}\left(E_i \mid A\right) = \frac{\mathbf{p}\left(A \mid E_i\right)\mathbf{p}\left(E_i\right)}{\sum_{j=1}^{n} \operatorname{dual}\mathbf{p}\left(A \mid E_j\right)\operatorname{dual}\mathbf{p}\left(E_j\right)} \qquad (3.1)$$

where E_i $(i = 1,\ldots,n)$ are mutually disjoint event partitions of Ω and $\sum_{j=1}^{n}\mathbf{p}\left(E_j\right) = 1$.

Based on the GIBR, the problem of cross-scale validation under variability and uncertainty can be formulated in several ways, such as single-point observation, multi-point observation, and multi-point multiscale observation.

5.1 Single-Point Observation

The simplest cross-scale validation is through the single-point observation. This approach allows that the uncertainty estimation at one scale is used to validate the model prediction at a different scale. Suppose the states of one or more variables $x_1,\cdots,x_l$ at Scale X are not directly observable. Instead, the system can be observed via one variable Y at Scale Y corresponding to the unobservable y at Scale Y. Then the estimation is calculated as follows.

Theorem 5.1. Given $\mathbf{p}\left(y \mid x_1,\ldots,x_L\right)$ for variables $x_1,\cdots,x_L$ at Scale X and y at Scale Y, $\mathbf{p}\left(Y \mid y\right)$ for observable Y corresponding to y, and the prior estimate $\mathbf{p}\left(x_1,\ldots,x_L\right)$, the posterior imprecise probability $\mathbf{p}\left(x_1,\ldots,x_L \mid Y\right)$ is obtained as

$$\mathbf{p}\left(x_1,\ldots,x_L \mid Y\right) = \frac{\mathbf{p}\left(x_1,\ldots,x_L\right)\int \mathbf{p}\left(Y \mid y\right)\mathbf{p}\left(y \mid x_1,\ldots,x_L\right)dy}{\operatorname{dual}\int\cdots\iint \mathbf{p}\left(Y \mid y\right)\mathbf{p}\left(y \mid x_1,\ldots,x_L\right)\mathbf{p}\left(x_1,\ldots,x_L\right)dydx_1\cdots dx_L}$$

5.2 An Illustrative Example of Cross-Scale Validation with Single-Point Observation

To illustrate the cross-scale validation with single-point observation, if we design an electrochemical biosensor from composites of carbon nanotubes (CNTs) and polymer and would like to know whether the electrical conductivity of the nanotube itself meets the design specification, experimental studies are needed. Instead of directly measuring the resistance of individual CNTs with diameters of about 10 nm, we can measure those from nanotube composites with the sizes of 1 μm or more, which is much easier and more accurate. The electrical conductivity of nanotubes is sensitively dependent on the geometry of tubes, particularly diameter and helicity. Because of variations of geometries and defects during the fabrication process, the measured quantities are stochastic in nature. At the same time, uncertainties are associated with measurements because of errors.

First, with the probability $\mathbf{p}\left(Y \mid y\right)$, the distribution of multiple measurements of Y reflects the true but unknown value y which can be estimated from prior experiences. The impreciseness of $\mathbf{p}\left(Y \mid y\right)$ is due to measurement errors. Second, there is a relationship between the large scale conductivity y of the composite (including both polymer and carbon nanotubes) and the small scale resistivity of x for individual CNTs. Therefore, the distribution $\mathbf{p}\left(y \mid x\right)$ corresponding to different possible values of y and x characterizes the probabilistic relationship between the large and small scale conductivities. The distribution $\mathbf{p}\left(y \mid x\right)$ can be estimated with available data, or by expert's experiences and judgements. Lack of precise data and information is the major source of uncertainty. Furthermore, we may or may not have prior estimates of $\mathbf{p}\left(x\right)$. If there is no data available, the complete ignorance of $\mathbf{p}\left(x = a\right) = [0,1]$ for any values of a can be applied. Then, based on the GIBR in Definition 5, we can update the information or the belief of the small scale conductivity distribution based on the new observation of the large scale conductivity $Y = e$ from an experiment by

$$\mathbf{p}\left(x = a \mid Y = e\right) = \frac{\mathbf{p}\left(x = a\right)\sum_j\left[\mathbf{p}\left(Y = e \mid y = c_j\right)\mathbf{p}\left(y = c_j \mid x = a\right)\right]}{\sum_i\sum_j\left[\begin{array}{l}\operatorname{dual}\mathbf{p}\left(Y = e \mid y = c_j\right)\\ \times\operatorname{dual}\mathbf{p}\left(y = c_j \mid x = a_i\right)\operatorname{dual}\mathbf{p}\left(x = a_i\right)\end{array}\right]}$$

This information update process can continue iteratively with more measurements Y's until $\mathbf{p}\left(x \mid Y\right)$'s converge to a satisfactory distribution with reduced uncertainty. That is, the widths of imprecise probabilities are reduced towards the precise ones.

Table 2 lists two sets of samples that measure the resistance of individual CNTs published in literature. Notice that the first paper as shown in the left column reported measurement errors or uncertainties with the ± ranges. The second paper did not record ranges. However, the first and fifth samples are right-censored and recorded with "≥" sign. The imprecise and incomplete information is the source of uncertainties. It is obvious that these two sets of data are inconsistence, which also shows the importance of imprecise probability in such applications.

We build probability distributions from the data in the first column of Table 2. The empirical cumulative distribution function (CDF) for each of the lower, middle, and upper observations are shown in Figure 2. They are solid lines in the colors of green, red, and blue respectively. If a parametric distribution is required, we can also fit the data by the Lognormal distributions, plotted as dotted curves, for three sets of data in terms of maximum likelihood. The estimated parameters for three distributions are:

- lower bound: $\mu_l = 3.4251, \sigma_l = 0.920699$
- middle: $\mu = 3.52202, \sigma = 0.920571$
- upper bound: $\mu_u = 3.6083, \sigma_u = 0.922064$

There are several ways to select the values of parameters for interval probabilities, particularly for multi-parameter distributions [44,45]. To simplify the illustration, we choose a simple way with the interval $[\mu_l, \mu_u] = [3.4251, 3.6083]$ and real-valued $\sigma = 0.920571$. Therefore, instead of using one single precise distribution to describe the distribution of the unit-length resistance of individual CNTs, we use the interval probability density function (PDF)

$$\mathbf{f}(x) = \frac{1}{\sqrt{2\pi}x\sigma} e^{-\frac{\left(\ln x - [\mu_l, \mu_u]\right)^2}{2\sigma^2}} \qquad (3.2)$$

Given the limited number of data, it is risky to use the empirical distributions or the fitted interval parametric models directly to represent the possible variations. A more cautious way is to build the lower and upper distributions based on the Kolmogorov-Smirnov confidence limit [31]. The 95% confidence lower and upper limits from the middle observation (red line) are shown as the blue and green dashed lines respectively in Figure 2. They are calculated by $\min\left(1, \max\left(0, \rho \pm D_{\alpha,n}\right)\right)$ where $D_{\alpha,n}$ depends on the sample size n and confidence level α. Here, n = 6, α = 0.025, and $D_{0.025,6}$ = 0.51926. This confidence band ensures that the probability of the unknown distribution function being within the band is at least 95%.

With the p-box formed by the Kolmogorov-Smirnov confidence limits and the Dempster-Shafer's structure of basic probability assignment (BPA) $m: 2^{\mathcal{A}} \to [0,1]$, we can determine the lower and upper probabilities of the resistivity. Specifically, the p-box is viewed as a stack of rectangles. The width of each rectangle is the focal element that defines the interval range of a BPA, whereas the height of the rectangle is the value of the BPA. The BPAs are:

$$m\left(0 \le \rho < 46.0\right) = 3/6 - D_{0.025,6} = 0.1474,$$

$$m\left(0 \le \rho < 48.9\right) = 1/6,$$

$$m\left(0 \le \rho < 117\right) = 1/6,$$

$$m\left(0 \le \rho < \infty\right) = D_{0.025,6} - \left(1 - D_{0.025,6}\right) = 0.0385,$$

$$m\left(7.8 \le \rho < \infty\right) = 1/6,$$

$$m\left(19.5 \le \rho < \infty\right) = 1/6,$$

$$m\left(37.6 \le \rho < \infty\right) = 1 - 2/6 - D_{0.025,6} = 0.1474.$$

Based on the Dempster-Shafer's belief function

$$\underline{p}\left(A\right) = \sum_{i:A_i \subseteq A} m\left(A_i\right)$$

and plausibility function

$$\overline{p}\left(A\right) = \sum_{i:A_i \cap A \neq \varnothing} m\left(A_i\right)$$

we can find the lower and upper probabilities. For instance, the lower and upper probabilities that the individual CNT resistivity is less than 50 Ω·m are $\underline{p}\left(\rho < 50\right) = m\left(0 \le \rho < 46.0\right) + m\left(0 \le \rho < 48.9\right) = 0.3140$ and $\overline{p}\left(\rho < 50\right) = 1$ respectively.

Table 2: Resistivity measurements of individual carbon nanotubes

Resistivity (Ω·m) [46] $\rho \pm \delta$ — *6 samples*	Resistivity (Ω·m) [47] ρ — *8 samples*
19.5±2.0	≥ 80
7.8±1.0	0.012
46.0±1.8	0.0075
37.6±1.0	580
48.9±4.3	≥ 0.4
117±19	0.00051
	0.098
	0.020

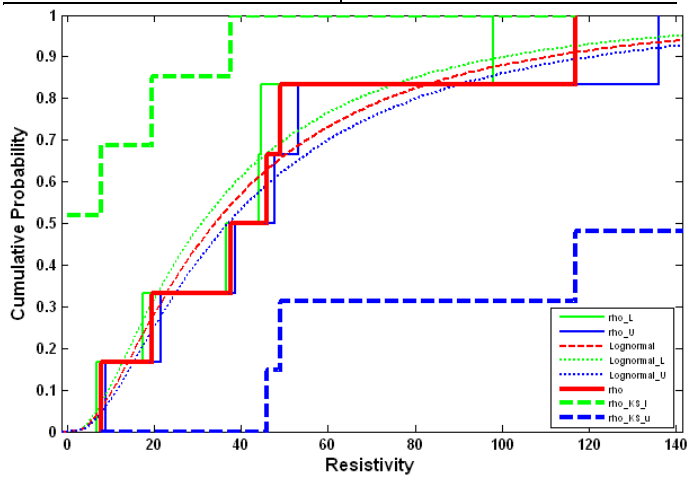

Figure 2: Empirical CDFs and the fitted distributions of data from Ref. [46]

Compared to individual CNT measurement, the measurement for CNT polymer composite is much easier to achieve. More than 200 publications have reported on the electrical property of CNT polymer composite. Bauhofer and Kovacs [48] recently summarized those experimental results, as shown in Figure 3. From the data in Ref. [48], the conductivity of composite with CNT concentration 1.0 wt% is compiled and listed in Table 3. The empirical CDF is plotted in Figure 4 as the red line. In addition, the 95% Kolmogorov-Smirnov confidence limits (dash lines) are calculated as lower and upper probability bounds. Here, $D_{0.025,17} = 0.31796$.

Similar to the above, the lower and upper probabilities can be determined based on the belief and plausibility functions. For instance, the lower and upper probabilities that the conductivity of CNT composites with 1.0 wt% is less than 0.1 are $\underline{p}(\sigma < 0.1) = 0.1526$ and $\overline{p}(\sigma < 0.1) = 0.6121$.

To simplify notation, we denote $\mathbf{p}(x) = \mathbf{p}(\rho < 50) = [0.3140, 1]$. Then $\mathbf{p}(x^C) = \mathbf{p}(\rho \geq 50) = 1 - \text{dual}\,\mathbf{p}(x) = [0.6860, 0]$. Assume that $\mathbf{p}(y \mid x) = \mathbf{p}(\sigma < 0.1) = [0.1526, 0.6121]$, which is the probability of the CNT composite conductivity is less than 0.1 $\Omega^{-1}\cdot m^{-1}$ if the individual CNT resistivity is less than 50 $\Omega\cdot m$. For instance, if the CNTs are all supplied by one vendor with good quality control, the resistivity of CNTs is known to be in a range. $\mathbf{p}(y^C \mid x) = 1 - \text{dual}\,\mathbf{p}(y \mid x) = [0.8474, 0.3879]$. No assumption is made about the probability of composite conductivity if the individual CNT resistivity is greater than 50 $\Omega\cdot m$. That is, $\mathbf{p}(y \mid x^C) = [0, 1]$, and it represents the total ignorance.

As a result, $\mathbf{p}(y^C \mid x^C) = 1 - \text{dual}\,\mathbf{p}(y \mid x^C) = [1, 0]$. Further, we assume the measurement is fairly reliable with $\mathbf{p}(Y \mid y) = [0.8, 0.9]$ and $\mathbf{p}(Y^C \mid y^C) = [0.8, 0.9]$. Thus $\mathbf{p}(Y^C \mid y) = [0.2, 0.1]$ and $\mathbf{p}(Y \mid y^C) = [0.2, 0.1]$. Now, if an additional experimental observation of Y that $\sigma < 0.1$ is obtained, then based on Theorem 5.1 we can assert that

$$\mathbf{p}(x \mid Y) = \frac{\mathbf{p}(x)\left[\mathbf{p}(Y \mid y)\mathbf{p}(y \mid x) + \mathbf{p}(Y \mid y^C)\mathbf{p}(y^C \mid x)\right]}{\text{dual}\begin{bmatrix}\mathbf{p}(Y \mid y)\mathbf{p}(y \mid x)\mathbf{p}(x) \\ +\mathbf{p}(Y \mid y)\mathbf{p}(y \mid x^C)\mathbf{p}(x^C) \\ +\mathbf{p}(Y \mid y^C)\mathbf{p}(y^C \mid x)\mathbf{p}(x) \\ +\mathbf{p}(Y \mid y^C)\mathbf{p}(y^C \mid x^C)\mathbf{p}(x^C)\end{bmatrix}} = [0.4002, 1]$$

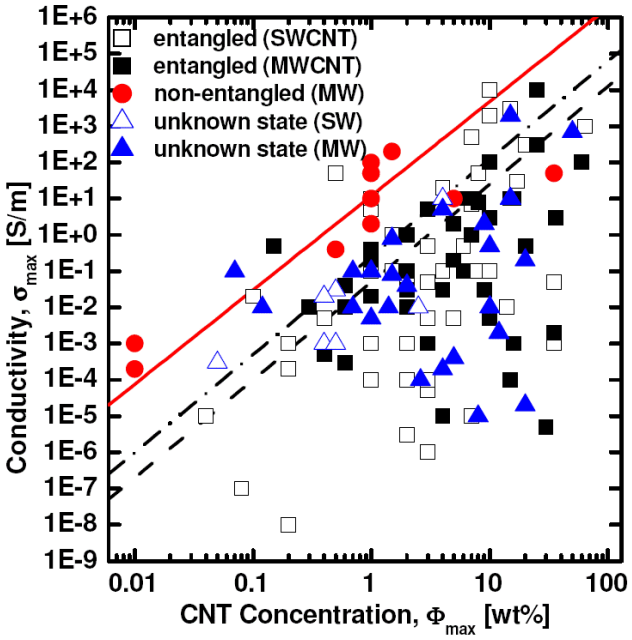

Figure 3: Conductivity measurements of CNT composites [48]

Table 3: Conductivity measurements of CNT polymer composites with CNT concentration of 1.0wt%

Maximum conductivity σ ($\Omega^{-1}\cdot m^{-1}$)	Number of Samples	Maximum conductivity σ ($\Omega^{-1}\cdot m^{-1}$)	Number of Samples
1.0×10^{-4}	1	4.0×10^{-1}	1
1.0×10^{-3}	1	2.0	1
5.0×10^{-3}	1	5.0	1
2.0×10^{-2}	2	1.0×10^{1}	2
1.0×10^{-1}	3	5.0×10^{1}	1
2.0×10^{-1}	1	1.0×10^{2}	1
3.0×10^{-1}	1		

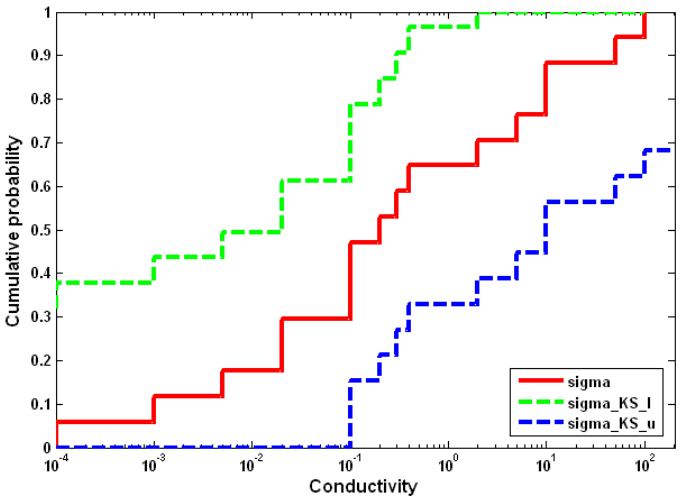

Figure 4: Empirical CDF of composites conductivity with 1.0 wt% of CNT from Ref. [48]

This posterior probability shows that the uncertainty level of individual CNT property is reduced to $1-0.4002=0.5998$ from the prior estimate of $1-0.3140=0.6860$.

Notice that interval probabilities allow us to calculate posterior probabilities even no data are available. When the total ignorance of $\mathbf{p}=[0,1]$ is applied, there is no risk of assuming certain prior probabilities, which is usually required in the traditional precise probability. Therefore, the calculated imprecise posterior probabilities are much more robust than the ones from the traditional Bayesian analysis with precise probabilities. In addition, the calculation of interval posterior probabilities based on our definition of imprecise probability has a much simpler form than other forms of imprecise probabilities, where linear or nonlinear optimization processes are required. This shows the significant advantage of our imprecise probability with the generalized interval form. Simple algebraic calculation is only required in computing probability intervals.

In summary, the example in this section demonstrates that the observation or measurement at one scale can be used to quantify and assess the variability and uncertainty of a relevant quantity at a different scale. This process can be useful to validate models or hypotheses that are concerned with quantities that are difficult or costly to measure. Instead, measurements at a different scale can still be applied to validate if intrinsic relations or dependencies between quantities of two scales exist.

5.3 Multi-Point Observation

Compared to the single-point observation approach, a more efficient approach to reduce the uncertainty associated with the estimate $\mathbf{p}\left(x_1,\ldots,x_l\right)$ is the multi-point observation. If there are multiple points of observation $Y_1,\ldots,Y_m$, Then the estimates of $x_1,\ldots,x_l$ can be more accurate.

Theorem 5.2. Given $\mathbf{p}\left(y_1,\ldots,y_M \mid x_1,\ldots,x_L\right)$ for variables $x_1,\cdots,x_L$ at Scale X and $y_1,\ldots,y_M$ at Scale Y, $\mathbf{p}\left(Y_1,\ldots,Y_M \mid y_1,\ldots,y_M\right)$ for observable Y_m's corresponding to y_m's ($m=1,\ldots,M$), and the prior estimate $\mathbf{p}\left(x_1,\ldots,x_L\right)$, the posterior imprecise probability $\mathbf{p}\left(x_1,\ldots,x_L \mid Y_1,\ldots,Y_M\right)$ is obtained as

$$\mathbf{p}\left(x_1,\ldots,x_L \mid Y_1,\ldots,Y_M\right) = \frac{\mathbf{p}\left(x_1,\ldots,x_L\right)\int\cdots\int\left[\begin{array}{l}\prod_{m=1}^{M}\mathbf{p}\left(Y_m \mid y_m\right)\\ \times\prod_{m=1}^{M}\mathbf{p}\left(y_m \mid x_1,\ldots,x_L\right)\end{array}\right]dy_1\cdots dy_M}{\operatorname{dual}\int\cdots\int\left[\begin{array}{l}\prod_{m=1}^{M}\mathbf{p}\left(Y_m \mid y_m\right)\\ \times\prod_{m=1}^{M}\mathbf{p}\left(y_m \mid x_1,\ldots,x_L\right)\\ \times\mathbf{p}\left(x_1,\ldots,x_L\right)\end{array}\right]dy_1\cdots dy_M dx_1\cdots dx_L}$$

5.4 An Illustrative Example of Cross-Scale Validation with Multi-Point Observation

We still use the CNT composite example in Section 5.2 to illustrate. The conductivity of the composite material is correlated with the concentration of CNT or the ratio of weights between CNT and polymer. The general trend observed in Figure 3 is that more CNT leads to the higher conductivity. Therefore, if the conductivities of two composites with different CNT concentrations, e.g. 1.0 wt% and 2.0 wt%, are measured as $Y_1=d_1$ and $Y_2=d_2$ respectively. The estimate $\mathbf{p}\left(x\right)$ can be updated by

$$\mathbf{p}\left(x=a \mid Y_1=e, Y_2=f\right) = \frac{\mathbf{p}\left(x=a\right)\sum_{j_1}\sum_{j_2}\left[\begin{array}{l}\mathbf{p}\left(Y_1=e \mid y_1=c_{j_1}\right)\\ \times\mathbf{p}\left(Y_2=f \mid y_2=d_{j_2}\right)\\ \times\mathbf{p}\left(y_1=c_{j_1} \mid x=a\right)\\ \times\mathbf{p}\left(y_2=d_{j_2} \mid x=a\right)\end{array}\right]}{\sum_i\sum_{j_1}\sum_{j_2}\left[\begin{array}{l}\operatorname{dual}\mathbf{p}\left(Y_1=e \mid y_1=c_{j_1}\right)\\ \times\operatorname{dual}\mathbf{p}\left(Y_2=f \mid y_2=d_{j_2}\right)\\ \times\operatorname{dual}\mathbf{p}\left(y_1=c_{j_1} \mid x=a_i\right)\\ \times\operatorname{dual}\mathbf{p}\left(y_2=d_{j_2} \mid x=a_i\right)\\ \times\operatorname{dual}\mathbf{p}\left(x=a_i\right)\end{array}\right]}$$

The information update process can continue with multiple measurements of Y_1 and Y_2.

From the data in Ref. [48], the conductivity of composite with CNT concentration 2.0 wt% is compiled and listed in Table 4. The CDFs are plotted in Figure 5. The lower and upper probabilities that $\sigma<0.1$ are $\underline{p}\left(\sigma<0.1\right)=0.5412$ and $\overline{p}\left(\sigma<0.1\right)=1$.

With $\mathbf{p}\left(x\right)=\left[0.3140,1\right]$, $\mathbf{p}\left(y_1 \mid x\right)=\left[0.1526,0.6121\right]$, $\mathbf{p}\left(y_2 \mid x\right)=\left[0.5412,1\right]$, $\mathbf{p}\left(y_1 \mid x^C\right)=\left[0,1\right]$, $\mathbf{p}\left(y_2 \mid x^C\right)=\left[0,1\right]$, $\mathbf{p}\left(Y_i \mid y_i\right)=\left[0.8,0.9\right]$, and

$\mathbf{p}\left(Y_i^C \mid y_i^C\right)=\left[0.8,0.9\right]$ for $i=1,2$, we can find

$$\mathbf{p}\left(x \mid Y_1,Y_2\right)=$$

$$\frac{\mathbf{p}(x)\begin{bmatrix}\mathbf{p}\left(Y_1 \mid y_1\right)\mathbf{p}\left(Y_2 \mid y_2\right)\mathbf{p}\left(y_1 \mid x\right)\mathbf{p}\left(y_2 \mid x\right)\\ +\mathbf{p}\left(Y_1 \mid y_1^C\right)\mathbf{p}\left(Y_2 \mid y_2\right)\mathbf{p}\left(y_1^C \mid x\right)\mathbf{p}\left(y_2 \mid x\right)\\ +\mathbf{p}\left(Y_1 \mid y_1\right)\mathbf{p}\left(Y_2 \mid y_2^C\right)\mathbf{p}\left(y_1 \mid x\right)\mathbf{p}\left(y_2^C \mid x\right)\\ +\mathbf{p}\left(Y_1 \mid y_1^C\right)\mathbf{p}\left(Y_2 \mid y_2^C\right)\mathbf{p}\left(y_1^C \mid x\right)\mathbf{p}\left(y_2^C \mid x\right)\end{bmatrix}}{\operatorname{dual}\begin{bmatrix}\mathbf{p}\left(Y_1 \mid y_1\right)\mathbf{p}\left(Y_2 \mid y_2\right)\mathbf{p}\left(y_1 \mid x\right)\mathbf{p}\left(y_2 \mid x\right)\mathbf{p}(x)\\ +\mathbf{p}\left(Y_1 \mid y_1^C\right)\mathbf{p}\left(Y_2 \mid y_2\right)\mathbf{p}\left(y_1^C \mid x\right)\mathbf{p}\left(y_2 \mid x\right)\mathbf{p}(x)\\ +\mathbf{p}\left(Y_1 \mid y_1\right)\mathbf{p}\left(Y_2 \mid y_2^C\right)\mathbf{p}\left(y_1 \mid x\right)\mathbf{p}\left(y_2^C \mid x\right)\mathbf{p}(x)\\ +\mathbf{p}\left(Y_1 \mid y_1^C\right)\mathbf{p}\left(Y_2 \mid y_2^C\right)\mathbf{p}\left(y_1^C \mid x\right)\mathbf{p}\left(y_2^C \mid x\right)\mathbf{p}(x)\\ +\mathbf{p}\left(Y_1 \mid y_1\right)\mathbf{p}\left(Y_2 \mid y_2\right)\mathbf{p}\left(y_1 \mid x^C\right)\mathbf{p}\left(y_2 \mid x^C\right)\mathbf{p}\left(x^C\right)\\ +\mathbf{p}\left(Y_1 \mid y_1^C\right)\mathbf{p}\left(Y_2 \mid y_2\right)\mathbf{p}\left(y_1^C \mid x^C\right)\mathbf{p}\left(y_2 \mid x^C\right)\mathbf{p}\left(x^C\right)\\ +\mathbf{p}\left(Y_1 \mid y_1\right)\mathbf{p}\left(Y_2 \mid y_2^C\right)\mathbf{p}\left(y_1 \mid x^C\right)\mathbf{p}\left(y_2^C \mid x^C\right)\mathbf{p}\left(x^C\right)\\ +\mathbf{p}\left(Y_1 \mid y_1^C\right)\mathbf{p}\left(Y_2 \mid y_2^C\right)\mathbf{p}\left(y_1^C \mid x^C\right)\mathbf{p}\left(y_2^C \mid x^C\right)\mathbf{p}\left(x^C\right)\end{bmatrix}}$$

$$=\left[0.6364,1\right]$$

Table 4: Conductivity measurements of CNT polymer composites with CNT concentration of 2.0wt%

Maximum conductivity σ ($\Omega^{-1}\cdot m^{-1}$)	Number of Samples	Maximum conductivity σ ($\Omega^{-1}\cdot m^{-1}$)	Number of Samples
3.0×10^{-6}	1	4.0×10^{-2}	1
1.0×10^{-4}	1	5.0×10^{-2}	2
1.0×10^{-3}	1	1.0×10^{-1}	2
1.0×10^{-2}	2	1.0	1
3.0×10^{-2}	1		

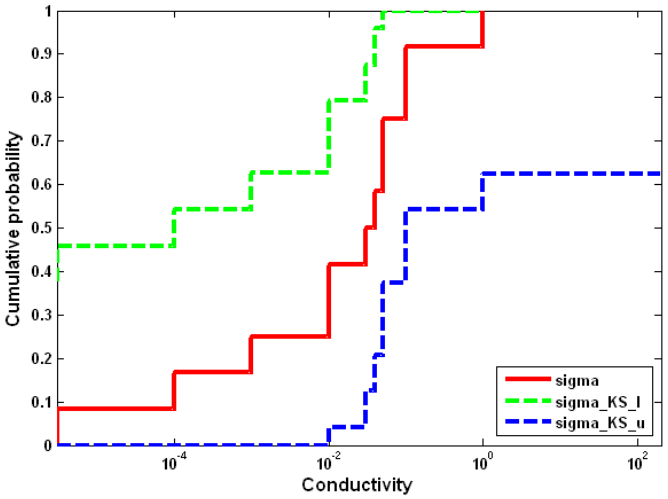

Figure 5: Empirical CDF of composites conductivity with 2.0 wt% of CNT from Ref. [48]

Compared to the single-point observation in Section 5.2, the uncertainty level is reduced faster, since more information of observation is used to assess the small scale quantity. Multi-point observation is an enhancement of the single-point observation and provides more information. Therefore, the uncertainty reduction and convergence to the satisfactory distribution should also be faster.

5.5 *Multi-Point Multiscale Observation*

As a further extension of the multi-point observation approach, the experimental measures can be conducted at two or more scales to validate the prediction of a model at a scale that is not observable. The observations from multiple scales and multiple sources can provide much more information about a system than single-scale observations. The information fusion of the multiscale observations can significantly reduce the uncertainty levels. As a result, multi-point multiscale observation is ideal to study variability and uncertainty of multiscale systems. The multiscale information fusion process can be conducted based on the following theorem.

Theorem 5.3. Given $\mathbf{p}\left(y_1,\ldots,y_M \mid x_1,\ldots,x_L\right)$ for variables $x_1,\cdots,x_L$ at Scale X and $y_1,\ldots,y_M$ at Scale Y, $\mathbf{p}\left(z_1,\ldots,z_N \mid y_1,\ldots,y_M\right)$ for variables $y_1,\ldots,y_M$ at Scale Y and $z_1,\cdots,z_N$ at Scale Z, $\mathbf{p}\left(Y_1,\ldots,Y_m \mid y_1,\ldots,y_m\right)$ for observable Y_m's corresponding to y_m's ($m=1,\ldots,M$), $\mathbf{p}\left(Z_1,\ldots,Z_n \mid z_1,\ldots,z_n\right)$ for observable Z_n's corresponding to z_n's ($n=1,\ldots,N$), and the prior estimate $\mathbf{p}\left(x_1,\ldots,x_L\right)$, the posterior imprecise probability $\mathbf{p}\left(x_1,\ldots,x_L \mid Y_1,\ldots,Y_M,Z_1,\ldots,Z_N\right)$ is obtained as

$$\mathbf{p}\left(x_1,\ldots,x_L \mid Y_1,\ldots,Y_M,Z_1,\ldots,Z_N\right)=$$

$$\frac{\mathbf{p}\left(x_1,\ldots,x_L\right)\int\cdots\int\begin{bmatrix}\mathbf{p}\left(Z_1,\ldots,Z_N \mid z_1,\ldots,z_N\right)\\ \mathbf{p}\left(Y_1,\ldots,Y_M \mid y_1,\ldots,y_M\right)\\ \mathbf{p}\left(z_1,\ldots,z_N \mid y_1,\ldots,y_M\right)\\ \mathbf{p}\left(y_1,\ldots,y_M \mid x_1,\ldots,x_L\right)\end{bmatrix}dz_1\cdots dz_N dy_1\cdots dy_M}{\operatorname{dual}\int\cdots\int\begin{bmatrix}\mathbf{p}\left(Z_1,\ldots,Z_N \mid z_1,\ldots,z_N\right)\\ \mathbf{p}\left(Y_1,\ldots,Y_M \mid y_1,\ldots,y_M\right)\\ \mathbf{p}\left(z_1,\ldots,z_N \mid y_1,\ldots,y_M\right)\\ \mathbf{p}\left(y_1,\ldots,y_M \mid x_1,\ldots,x_L\right)\\ \mathbf{p}\left(x_1,\ldots,x_L\right)\end{bmatrix}dz_1\cdots dz_N dy_1\cdots dy_M dx_1\cdots dx_L}$$

6. SUMMARY

In this paper, a generalized multiscale Markov model is proposed to represent variability and uncertainty simultaneously in analyzing multiscale systems. The model

captures uncertainty propagation across different length scales to support complex system analysis.

The proposed model is based on a new theory of imprecise probability that has the new form of generalized interval, where proper and improper intervals capture uncertainties. With an algebraic structure similar to the precise probability, the new definition of imprecise probability significantly simplifies the inference and reasoning compared to other forms of imprecise probabilities. The precise probability becomes just a special case of the proposed one, where the widths of probabilities are reduced to zeros. Three graphoid properties of independence based on the new form of imprecise probability are studied, as independence is essential to Markov models. The proposed multiscale Markov model allows us to compute the propagation of stochastic and uncertain information across length scales efficiently. This is enabled by a new definition of Bayes' rule with the generalized imprecise probability. Three cross-scale validation approaches are formulated so that information fusion can be achieved.

The proposed model and inference mechanisms are generic in nature. The proposed model supports the design and analysis of any multiscale system when variability and uncertainty are considered. The simplicity of the reasoning based on the proposed model shows the advantages and potentials for wider applications in science and engineering, compared to other forms of imprecise probabilities. The future extension includes the study of simulation approaches based on the proposed model and their applications to various multiscale system design problems. Further investigation on the fundamental properties of the new imprecise probability is also needed.

APPENDICES

A. Proof of Theorem 3.1

Proof. $\mathbf{p}\left(A \cap B \mid C\right) = \mathbf{p}\left(A \mid C\right)\mathbf{p}\left(B \mid C\right) \Leftrightarrow$

$\mathbf{p}\left(A \cap B \cap C\right) / \operatorname{dual}\mathbf{p}\left(C\right) = \mathbf{p}\left(A \mid C\right) \cdot \mathbf{p}\left(B \cap C\right) / \operatorname{dual}\mathbf{p}\left(C\right)$

$\Leftrightarrow \quad \mathbf{p}\left(A \cap B \cap C\right) / \operatorname{dual}\mathbf{p}\left(B \cap C\right) = \mathbf{p}\left(A \mid C\right) \Leftrightarrow$

$\mathbf{p}\left(A \mid B \cap C\right) = \mathbf{p}\left(A \mid C\right)$.

□

B. Proof of Corollary 3.2

Proof. $X \perp Y \mid Z \Rightarrow \mathbf{p}\left(X \cap Y \mid Z\right) = \mathbf{p}\left(X \mid Z\right)\mathbf{p}\left(Y \mid Z\right)$

$\Rightarrow \mathbf{p}\left(Y \cap X \mid Z\right) = \mathbf{p}\left(Y \mid Z\right)\mathbf{p}\left(X \mid Z\right) \Rightarrow Y \perp X \mid Z$.

□

C. Proof of Corollary 3.3

Proof. $\mathbf{p}\left(X \mid Y \cap X\right) = \mathbf{p}\left(X \mid Y\right)$.

□

D. Proof of Corollary 3.4

Proof. $X \perp W \mid \left(Y \cap Z\right) \Rightarrow$

$\mathbf{p}\left(X \mid W \cap \left(Y \cap Z\right)\right) = \mathbf{p}\left(X \mid Y \cap Z\right)$. Further, $X \perp Y \mid Z$

$\Rightarrow \quad \mathbf{p}\left(X \mid Y \cap Z\right) = \mathbf{p}\left(X \mid Z\right)$. Therefore,

$\mathbf{p}\left(X \mid \left(W \cap Y\right) \cap Z\right) = \mathbf{p}\left(X \mid Z\right) \Rightarrow X \perp \left(W \cap Y\right) \mid Z$.

□

E. Proof of Theorem 4.1

Proof. By the definitions of conditional probability in Eq.(2.3) and independence in Eq.(2.5), we have

$$\mathbf{p}\left(X_{\mathcal{A}_i}, X_{\mathcal{A}_j} \mid x_{i\in\mathcal{A}_i}, x_{j\in\mathcal{A}_j}\right) = \frac{\mathbf{p}\left(X_{\mathcal{A}_i}, X_{\mathcal{A}_j}, x_{i\in\mathcal{A}_i}, x_{j\in\mathcal{A}_j}\right)}{\operatorname{dual}\mathbf{p}\left(x_{i\in\mathcal{A}_i}, x_{j\in\mathcal{A}_j}\right)}$$

$$= \frac{\mathbf{p}\left(X_{\mathcal{A}_i}, x_{i\in\mathcal{A}_i}\right)\mathbf{p}\left(X_{\mathcal{A}_j}, x_{j\in\mathcal{A}_j}\right)}{\operatorname{dual}\left[\mathbf{p}\left(x_{i\in\mathcal{A}_i}\right)\mathbf{p}\left(x_{j\in\mathcal{A}_j}\right)\right]} = \frac{\mathbf{p}\left(X_{\mathcal{A}_i}, x_{i\in\mathcal{A}_i}\right)\mathbf{p}\left(X_{\mathcal{A}_j}, x_{j\in\mathcal{A}_j}\right)}{\operatorname{dual}\mathbf{p}\left(x_{i\in\mathcal{A}_i}\right)\operatorname{dual}\mathbf{p}\left(x_{j\in\mathcal{A}_j}\right)}$$

$$= \mathbf{p}\left(X_{\mathcal{A}_i} \mid x_{i\in\mathcal{A}_i}\right)\mathbf{p}\left(X_{\mathcal{A}_j} \mid x_{j\in\mathcal{A}_j}\right)$$

□

F. Proof of Theorem 4.2

Proof. By the definitions of conditional probability and independence, we have

$$\mathbf{p}\left(x_{\mathcal{A}_i} \mid y_{\mathcal{B}_j}, z_{\mathcal{C}_k}\right) = \frac{\mathbf{p}\left(x_{\mathcal{A}_i}, y_{\mathcal{B}_j}, z_{\mathcal{C}_k}\right)}{\operatorname{dual}\mathbf{p}\left(y_{\mathcal{B}_j}, z_{\mathcal{C}_k}\right)}$$

$$= \frac{\mathbf{p}\left(x_{\mathcal{A}_i}, y_{\mathcal{B}_j}\right)\mathbf{p}\left(z_{\mathcal{C}_k}\right)}{\operatorname{dual}\mathbf{p}\left(y_{\mathcal{B}_j}\right)\operatorname{dual}\mathbf{p}\left(z_{\mathcal{C}_k}\right)} = \mathbf{p}\left(x_{\mathcal{A}_i} \mid y_{\mathcal{B}_j}\right)$$

since $\mathbf{p}\left(z_{\mathcal{C}_k}\right) / \operatorname{dual}\mathbf{p}\left(z_{\mathcal{C}_k}\right) = 1$.

□

G. Proof of Theorem 5.1

Proof.

$$\mathbf{p}\left(x_1,\ldots,x_L \mid Y\right) = \frac{\mathbf{p}\left(Y, x_1,\ldots,x_L\right)}{\operatorname{dual}\mathbf{p}\left(Y\right)} = \frac{\int \mathbf{p}\left(Y, y, x_1,\ldots,x_L\right)dy}{\operatorname{dual}\mathbf{p}\left(Y\right)}$$

because of the logic coherent constraint $\int \mathbf{p}\left(y\right)dy = 1$.

$$\frac{\int \mathbf{p}\left(Y, y, x_1,\ldots,x_l\right)dy}{\operatorname{dual}\mathbf{p}\left(Y\right)}$$

$$= \frac{\mathbf{p}\left(x_1,\ldots,x_l\right)\int \mathbf{p}\left(Y \mid y\right)\mathbf{p}\left(y \mid x_1,\ldots,x_l\right)dy}{\operatorname{dual}\int\cdots\iint \mathbf{p}\left(Y \mid y\right)\mathbf{p}\left(y \mid x_1,\ldots,x_l\right)\mathbf{p}\left(x_1,\ldots,x_l\right)dydx_1\cdots dx_l}$$

because of Theorems 3.1.

□

H. Proof of Theorem 5.2

Proof.

$$\mathbf{p}\left(x_1,\ldots,x_L \mid Y_1,\ldots,Y_M\right)$$

$$= \frac{\mathbf{p}\left(Y_1,\ldots,Y_M, x_1,\ldots,x_L\right)}{\operatorname{dual}\mathbf{p}\left(Y_1,\ldots,Y_M\right)}$$

$$= \frac{\int\cdots\int \mathbf{p}\left(Y_1,\ldots,Y_M,y_1,\ldots,y_M,x_1,\ldots,x_L\right)dy_1\cdots dy_M}{\operatorname{dual}\int\cdots\int \mathbf{p}\left(Y_1,\ldots,Y_M,y_1,\ldots,y_M,x_1,\ldots,x_L\right)dy_1\cdots dy_M dx_1\cdots dx_L}$$

$$= \frac{\int\cdots\int \begin{bmatrix}\mathbf{p}\left(Y_1,\ldots,Y_M \mid y_1,\ldots,y_M\right)\\ \times\mathbf{p}\left(y_1,\ldots,y_M \mid x_1,\ldots,x_L\right)\\ \times\mathbf{p}\left(x_1,\ldots,x_L\right)\end{bmatrix}dy_1\cdots dy_M}{\operatorname{dual}\int\cdots\int \begin{bmatrix}\mathbf{p}\left(Y_1,\ldots,Y_M \mid y_1,\ldots,y_M\right)\\ \times\mathbf{p}\left(y_1,\ldots,y_M \mid x_1,\ldots,x_L\right)\\ \times\mathbf{p}\left(x_1,\ldots,x_L\right)\end{bmatrix}dy_1\cdots dy_M dx_1\cdots dx_L}$$

because of Theorems 3.1. If $y_1,\ldots,y_M$ and their measurements $Y_1,\ldots,Y_M$ are mutually independent, from Theorem 4.1, the above can be simplified further to

$$= \frac{\mathbf{p}\left(x_1,\ldots,x_L \mid Y_1,\ldots,Y_M\right)\mathbf{p}\left(x_1,\ldots,x_L\right)\int\cdots\int\left[\prod_{m=1}^{M}\mathbf{p}\left(Y_m \mid y_m\right)\times\prod_{m=1}^{M}\mathbf{p}\left(y_m \mid x_1,\ldots,x_L\right)\right]dy_1\cdots dy_M}{\operatorname{dual}\int\cdots\int\left[\prod_{m=1}^{M}\mathbf{p}\left(Y_m \mid y_m\right)\times\prod_{m=1}^{M}\mathbf{p}\left(y_m \mid x_1,\ldots,x_L\right)\times\mathbf{p}\left(x_1,\ldots,x_L\right)\right]dy_1\cdots dy_M dx_1\cdots dx_L}$$

□

I. Proof of Theorem 5.3

Proof.

$$\mathbf{p}\left(x_1,\ldots,x_L \mid Y_1,\ldots,Y_M,Z_1,\ldots,Z_N\right) = \frac{\mathbf{p}\left(Z_1,\ldots,Z_N,Y_1,\ldots,Y_M,x_1,\ldots,x_L\right)}{\operatorname{dual}\mathbf{p}\left(Z_1,\ldots,Z_N,Y_1,\ldots,Y_M\right)}$$

$$= \frac{\int\cdots\int \mathbf{p}\left(Z_1,\ldots,Z_N,z_1,\ldots,z_N,Y_1,\ldots,Y_M,y_1,\ldots,y_M,x_1,\ldots,x_L\right)dz_1\cdots dz_N dy_1\cdots dy_M}{\operatorname{dual}\int\cdots\int \mathbf{p}\left(Z_1,\ldots,Z_N,z_1,\ldots,z_N,Y_1,\ldots,Y_M,y_1,\ldots,y_M,x_1,\ldots,x_L\right)dz_1\cdots dz_N dy_1\cdots dy_M dx_1\cdots dx_L}$$

$$= \frac{\mathbf{p}\left(x_1,\ldots,x_L\right)\int\cdots\int\left[\begin{matrix}\mathbf{p}\left(Z_1,\ldots,Z_N \mid z_1,\ldots,z_N\right)\mathbf{p}\left(Y_1,\ldots,Y_M \mid y_1,\ldots,y_M\right)\\ \mathbf{p}\left(z_1,\ldots,z_N \mid y_1,\ldots,y_M\right)\mathbf{p}\left(y_1,\ldots,y_M \mid x_1,\ldots,x_L\right)\end{matrix}\right]dz_1\cdots dz_N dy_1\cdots dy_M}{\operatorname{dual}\int\cdots\int\left[\begin{matrix}\mathbf{p}\left(Z_1,\ldots,Z_N \mid z_1,\ldots,z_N\right)\mathbf{p}\left(Y_1,\ldots,Y_M \mid y_1,\ldots,y_M\right)\\ \mathbf{p}\left(z_1,\ldots,z_N \mid y_1,\ldots,y_M\right)\mathbf{p}\left(y_1,\ldots,y_M \mid x_1,\ldots,x_L\right)\mathbf{p}\left(x_1,\ldots,x_L\right)\end{matrix}\right]dz_1\cdots dz_N dy_1\cdots dy_M dx_1\cdots dx_L}$$

□

REFERENCES

[1] Walley P. (1991) Statistical Reasoning with Imprecise Probabilities, Chapman & Hall, London.

[2] Draper D. (1995) Assessment and propagation of model uncertainty (with discussion). *Journal of the Royal Statistical Society*, **B57**(1): 45–97.

[3] Zimmermann H.J. (2000) An application-oriented view of modeling uncertainty. *European Journal of Operational Research*, **122**: 190-198

[4] Elishakoff I. and Ren Y. (2003) Finite Element Methods for Structures with Large Stochastic Variations, Oxford University Press

[5] Liu W.K., Belytschko T., and Mani A. (1986) Random field finite elements. *International Journal for Numerical Methods in Engineering*, **23**(10): 1831-1845

[6] Ghanem R.G. and Spanos P.D. (2003) *Stochastic Finite Elements: A Spectral Approach*, Dover Publications

[7] Xiu D. and Karniadakis G.E. (2003) Modeling uncertainty in flow simulations via generalized polynomial chaos. *Journal of Computational Physics*, **187**(1): 137-167

[8] Xiu D. and Hesthaven J.S. (2005) High-order collocation methods for differential equations with random inputs. *SIAM Journal on Scientific Computing*, **27**(3): 1118-1139

[9] Babuška I., Nobile F., and Tmpone R. (2007) A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, **45**(3): 1005-1034

[10] Sudret B. and Der Kiureghian A. (2000) Stochastic finite element methods and Reliability – A state-of-the-art report. University of California Berkeley, Report No. UCB/SEMM-2000/08

[11] Hähner P. (1996) A theory of dislocation cell formation based on stochastic dislocation dynamics. *Acta Materialia*, **44**(6): 2345-2352

[12] El-Azab A. (2000) Statistical mechanics treatment of the evolution of dislocation distributions in single crystals. *Physical Review B*, **61**(18): 11956-11966

[13] Zaiser M. (2001) Statistical modeling of dislocation systems. *Materials Science & Engineering A*, **309-310**: 304-315

[14] Hiratani M. and Zbib H.M. (2003) On dislocation-defect interactions and patterning: stochastic discrete dislocation

dynamics (SDD). *Journal of Nuclear Materials*, **323**(2-3): 290-303

[15] Liu W.K., Siad L., Tian R., Lee S., Lee D., Yin X., Chen W., Chan S. Olson G.B., Lindgen L.-E., Horstemeyer M.F., Chang Y.-S., Choi J.-B., and Kim Y.J. (2009) Complexity science of multiscale materials via stochastic computations. *Int. J. Numer. Meth. Engng*, **80**(6-7): 932-978

[16] Choi M.J., Chandrasekaran V., Malioutov D.M., Johnson J.K., and Willsky A.S. (2008) Multiscale stochastic modeling for tractable inference and data assimilation. *Computational Methods in Applied Mechanics & Engineering*, **197**(43-44): 3492-3515

[17] Chamoin L., Oden J.T., and Prudhomme S. (2008) A stochastic coupling method for atomic-to-continuum Monte-Carlo simulations. *Computational Methods in Applied Mechanics & Engineering*, **197**(43-44): 3530-3546

[18] Ganapathysubramanian B. and Zabaras N. (2008) Modeling multiscale diffusion processes in random heterogeneous media. *Computational Methods in Applied Mechanics & Engineering*, **197**(43-44): 3560-3573

[19] Arnst M. and Ghanem R. (2008) Probabilistic equivalence and stochastic model reduction in multiscale analysis. *Computational Methods in Applied Mechanics & Engineering*, **197**(43-44): 3584-3592

[20] Yin, X., Lee, S., Chen, W., Liu W.K., and Horstemeyer M.F. (2009) Efficient random field uncertainty propagation in design using multiscale analysis. *Journal of Mechanical Design*, **131**(2): 021006(1-10)

[21] Chen W., Yin X., Lee, S., and Liu W.K. (2009) A multiscale design methodology for designing hierarchical multiscale systems considering random field uncertainty. *Proc. 2009 ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conference (IDETC/CIE 2009)*, San Diego, CA, paper No. DETC2009-87465

[22] MacDonald I.L. and Zucchini W. (1997) Hidden Markov and Other Models for Discrete-Valued Time Series, Chapman & Hall

[23] Cappé O., Moulines E., and Rydén T. (2005) *Inference in Hidden Markov Models*, Springer

[24] Fine S., Singer Y., and Tishby N. (1998) The hierarchical hidden Markov model: Analysis and applications. *Machine Learning*, **32**: 41-62

[25] Bouman C.A. and Shapiro M. (1994) A multiscale random field model for Bayesian image segmentation. *IEEE Transactions on Image Processing*, **3**(2): 162-177

[26] Laferté J.M., Perez P., and Heitz F. (2000) Discrete Markov modeling and inference on the quad-tree. *IEEE Transactions on Image Processing*, **9**(3): 390-404

[27] Dempster A. (1967) Upper and lower probabilities induced by a multi-valued mapping. *Annals of Mathematical Statistics*, **38**(2):325—339

[28] Shafer G.A. (1990) *Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ.

[29] Dubois D. and Prade H. (1988) Possibility Theory: An Approach to Computerized Processing of Uncertainty. Plenum, New York.

[30] Molchanov I. (2005) *Theory of Random Sets*. London: Springer.

[31] Ferson S., Kreinovich V. Ginzburg L., Myers D.S., and Sentz K. (2003) Constructing probability boxes and Dempster-shafer structures. *Sandia National Laboratories Technical report* SAND2002-4015, Albuquerque, NM.

[32] Weichselberger K. (2000) The theory of interval-probability as a unifying concept for uncertainty. *International Journal of Approximate Reasoning*, **24**(2-3): 149-170

[33] Möller B. and Beer M. (2004) Fuzzy Randomness: Uncertainty in Civil Engineering and Computational Mechanics. Springer, Berlin.

[34] Neumaier A. (2004) Clouds, fuzzy sets, and probability intervals. *Reliable Computing*, **10**(4):249—272.

[35] Wang Y. (2008) Imprecise probabilities with a generalized interval form. In R.L. Muhanna and R.L. Mullen, eds., *Proc. 3rd Int. Workshop on Reliability Engineering Computing (REC'08)*, Savannah, Georgia, pp.45-59.

[36] Wang Y. (2010) Imprecise probabilities based on generalized intervals for system reliability assessment. To appear in *International Journal of Reliability & Safety*

[37] Kaucher E. (1980) Interval analysis in the extended interval space IR. *Computing Supplementa*, Vol.2, 33-49

[38] Moore R.E. (1966) *Interval Analysis*. Prentice-Hall, Englewood Cliffs, N.J.

[39] Gardeñes E., Sainz M.Á., Jorba L., Calm R., Estela R., Mielgo H., and Trepat A. (2001) Modal intervals. *Reliable Computing*, **7**(2): 77-111

[40] Dimitrova N.S., Markov S.M., and Popova E.D.(1992) Extended Interval Arithmetics: New Results and Applications. In L. Atanassova and J. Herzberger (Eds.) *Computer Arithmetic and Enclosure Methods*, pp.225-232

[41] Wang Y. (2008) Semantic Tolerance Modeling with Generalized Intervals. *Journal of Mechanical Design*, **130**(8): 081701(1-7)

[42] Wang Y. (2008) Closed-Loop Analysis in Semantic Tolerance Modeling. *Journal of Mechanical Design*, **130**(6): 061701(1-10)

[43] Wang Y. (2008) Interpretable Interval Constraint Solvers in Semantic Tolerance Analysis. *Computer-Aided Design & Applications*, **5**(5): 654-666

[44] Batarseh O.G. and Wang Y. (2008) Reliable simulation with input uncertainties using an interval-based approach. *Proc. 2008 Winter Simulation Conference*, Miami, Florida

[45] Batarseh O.G. (2010) An Interval-Based Approach to Model Input Uncertainty in Discrete-Event Simulation. Ph.D. dissertation, University of Central Florida.

[46] Dai H., Wong E.W., and Lieber C.M. (1996) Probing electrical transport in nanomaterials: conductivity of individual carbon nanotubes. *Science*, **272**(5261): 523-526

[47] Ebbesen T.W., Lezec H.J., Hiura H., Bennett J.W., Ghaemi H.F., and Thio T. (1996) Electrical conductivity of individual carbon nanotubes. *Nature*, **382**: 54-56
[48] Bauhofer W. and Kovacs J.Z. (2009) A review and analysis of electrical percolation in carbon nanotube polymer composites. *Composites Science & Technology*, **69**(10): 1486-1498