Virtual Model Validation of Complex Multiscale Systems: Applications to Nonlinear Elastostatics

J. Tinsley Oden*, Ernesto E. Prudencio and Paul T. Bauman

Institute for Computational Engineering and Sciences (ICES)
The University of Texas at Austin
201 East 24th St, Stop C0200, Austin, TX 78712-1229

Abstract

We propose a virtual statistical validation process as an aid to the design of experiments for the validation of phenomenological models of the behavior of material bodies, with focus on those cases in which knowledge of the fabrication process used to manufacture the body can provide information on the micro-molecular-scale properties underlying macroscale behavior. One example is given by models of elastomeric solids fabricated using polymerization processes. We describe a framework for model validation that involves Bayesian updates of parameters in statistical calibration and validation phases. The process enables the quantification of uncertainty in quantities of interest (QoIs) and the determination of model consistency using tools of statistical information theory. We assert that microscale information drawn from molecular models of the fabrication of the body provides a valuable source of prior information on parameters as well as a means for estimating model bias and designing virtual validation experiments to provide information gain over calibration posteriors.

Keywords: Model validation, information entropy, Kullback-Leibler divergence, mutual information, uncertainty quantification, Bayesian inference.

1. Introduction

The subject of validation of mathematical and computational models of physical events has become one of great interest in recent years as it is increasingly recognized to be the key to the reliability of all science-based predictions. There is a large and growing literature on the subject, including books and treatises [40, 57], papers and reports on basic definitions and principles (*e.g.* [1, 3, 4, 5, 6, 10, 21, 24, 28, 29, 41, 44, 45, 46]), examples of validation processes, (*e.g.* [11, 44, 47, 48]), statistical approaches to validation and the quantification of uncertainty in predictions (*e.g.* [10, 25]), and a comprehensive study undertaken by the U.S. National Research Council [17]. Still, there remain basic differences in the literature on how validation processes are defined, interpreted, and implemented; moreover, the development of effective methods for the validation of models of complex physical events, such as multiscale models of materials, remains a largely open problem. General Bayesian approaches to model validation with special attention to multiscale models of materials, are the subjects of this work.

*Corresponding author, oden@ices.utexas.edu

We have in mind here mainly developing approaches to assess the validity of models for predictions of events in the physical universe and in engineering applications and physical science, and particularly phenomenological models of continuum mechanics derived by averaging, in some sense, events at molecular or atomistic scales. We demonstrate that the use of such fine scale information can lead to important, often overlooked information of use in designing validation experiments and in assessing model bias. The use of fine-scale models to study macro-model behavior leads us to use the terms "virtual validation," a flagrant oxymoron. But in no way do we imply that such approaches are a substitute for validation experiments: they only provide a tremendously less expensive and a science-based guide to the design of effective experiments to aid in model validation processes.

There is also the implication that the fine scale model is somehow a superior representation of the underlying physics than the macro model, which may not be true. In many cases, such as those discussed here, the continuum model parameters can be deduced by averaging fine-scale information, but the validity of the fine-scale model may not be completely confirmed. This is a subject of current research (*e.g.* [20]). The characterization of a compatible molecular model that provides a basis for continuum predictions often represents a better characterization of our current knowledge about the physical setting of the problem and, therefore, should be used.

A few basic ideas are worth reviewing. We understand a mathematical model (of physical phenomena) to mean a collection of mathematical constructions (differential or algebraic equations, inequalities, etc.) designed to put in mathematical context both inductive hypotheses about the function and behavior of the physical system as well as observations of the response of the system to various inputs. Such models can be derived from general principles covering the physical phenomena, or a surrogate of a general theory produced by simplifying assumptions, or they can be empirical, produced by creating mathematical expressions that fit model outputs to features of experimental observations.

The word *valid*, according to the Merriam-Webster Collegiate Dictionary [36], means "well-grounded, justifiable, being at once relevant and meaningful, logically correct." The notion of validation is more forgiving and more in keeping with the fact that any measure of validity must cope with the inevitable uncertainties and incompleteness in data and observations. According, again, to Merriam-Webster, validation is a *process* of determining the degree of validity "(of a measuring device)." Thus, there are degrees of validation. A model can therefore attain a degree of validity according to some meaningful measure and that measure must take into account uncertainties in data and predictions. The Free Dictionary by Farlex, accessed through Google or Bing, provides as a definition of validation of computer models and simulation: *the process of determining the degree to which a model or simulation is an accurate representation of the real world from the perspective of the intended use of the model.* Again, a "degree" of accuracy is called for. This definition, precisely that found in the guide [3], is close to the definition in [40, 57]. In [45, 46], it is condensed to: *the process of determining the accuracy with which a model can predict specified features of physical reality.* We provide an alternative definition later in this section. Importantly, a single model may be used to predict several quantities of interest and the level of uncertainty of each may be quite different. It is worth noting that relatively few works on model validation actually address the degree of validation or how to measure it.

The principle of falsification of Popper [53] asserts that hypotheses cannot be accepted as scientific theory unless they can possibly be refuted by physical observations (and then, if in conflict with observations, the hypotheses are discarded or modified so as to be replaced by a new theory that is in line with observations). According to this principle, no theory can be, technically speaking, perfectly valid. By the same token, no mathematical model can be

perfectly valid if one accepts the idea that new observational data can be acquired that is in conflict with earlier predictions. But Popper's principle came under severe criticism as being too objective: ignoring the place in science of statistical theories and experimental uncertainties. So he modified the idea in [53] to cover statistical hypotheses not strictly falsefiable, but "practically falsefied" when the hypothesis attaches a sufficiently small probability to an event. But, according to Howson and Urbach [26], the objective mechanism by which a hypothesis may "be logically refuted by observational evidence could never work with statistical hypothesis, for these ascribe probabilities to possible events and do not say of any that they will or will not actually occur." But the goal of validation processes is not, strictly speaking, to find truth; it is to assess the accuracy of predictions and to determine if the accuracy is sufficient for an intended purpose. That the predictions and observational data may be probability distributions merely means one must develop measures of their relative accuracy and their uncertainty and, like Popper, establish tolerances that define their acceptability.

In the present paper, we describe a general approach to validation and UQ in a Bayesian setting; we fully subscribe to Jaynes' view [27] that the "logic of science", at least much of science, becomes transparent in a Bayesian setting. Beyond that, notions of information entropy [19, 59] and related ideas provide additional tools outside the Bayesian framework.

The subject of model validation is a central theme of what some refer to as "predictive science," the discipline concerned with the prediction of features of physical events or of the behavior of engineering systems. Predictions involve extrapolations based on available models and data from what is known and accessible to future events that may have not yet occurred. This begs the question of what qualifies as observational data and what are the actual goals of the prediction - the quantities of interest.

An observable is a physical entity that can be observed, that is, detected using the human senses, usually aided by instrumentation, usually corrupted by experimental error, and is the source of data in model calibration and validation processes.

Quantities of interest are specific features of the response of the system under study. They are the goals of the prediction that the model is called up upon to deliver. Processes of model validation are meaningful only when specific QoIs are defined: a model may be able to deliver a "valid" prediction (as defined in this paper) for one QoI, but not another.

QoIs are not observables. If an event is observed, that is, measured, it becomes, by definition, part of the validation data. The purpose of the use of models in predictions is to extrapolate beyond what is recorded as data so as to forecast events not yet observed. The ideas are at the heart of inductive logic and a fundamental notion of the scientific method. In fact, it has also been the foundational principle of engineering design from time immemorial. Engineering systems are designed to function in various ways under design conditions, but their actual performance cannot be observed before the system exists. Tests may be performed on sophisticated models of the final system, but these can only increase confidence in the likelihood of behavior predicted by the engineer - these are validation tests. The QoIs, the specific features of the response of the final system in the field, can be predicted but not observed. Once any measurement of a feature is taken, that measurement becomes observational data - it is no longer a quantity of interest as it is no longer a goal of the predictive effort - it is known to within experimental error.

Unfortunately, or perhaps, inevitably, this fact points to the subjective nature of science and the scientific method of prediction of future events, and, correspondingly, of the entire subject of model validation. This is why the Bayesian framework, described so eloquently in Jaynes [27], provides an ideal setting to discuss model validation. With these ideas in mind, we put forward a new notion of model validation as follows: *the process of determining the level of*

confidence one has in the ability of the model to predict quantities of interest to within preset tolerances based on the accuracy with which the model predicts specific observables.

We note that a significant body of literature exists that employs Bayesian approaches and information theoretics to address issues of model validation. The important work of Kennedy and O'Hagan [32, 33] called attention to the importance of model discrepancy or bias and its role in Bayesian validation processes; there the use of Gaussian processes (GPs) as a means to represent and calculate model bias and to accelerate computing model outputs in validation processes is advocated. Bayarri, *et al.*, [10] present a detailed analysis of a class of Gaussian processes for modeling model outputs and model bias, and discuss Bayesian methods for estimating model parameters and the so-called hyperparameters of GPs. The papers of Jaing and Mahaderan [28, 29] describe Bayesian-based validation processes that use Bayesian hypothesis testing and cross-entropy methodologies for the design of validation experiments. These latter approaches generally employ Kullback-Leibler divergence measures to compare successive probability distributions generated in sequences of validation experiments, the key characterizations being the "cross entropy" of pairs of distributions. Related work has been done by Foglia, *et al.* [22] and Weijs, *et al.* [64]. A comprehensive survey of cross-validation procedures for model selection has been presented by Arlot and Celisse [2].

Not all works on model validation advocate Bayesian approaches. In Ferson, *et al.* [21], the use of Bayesian approaches is said to be "not the proper focus on validation," as interest should be "in objectively measuring the conformances of predictions against data that have not been used previously to calibrate the model." They go on to say that "when predictions are distributions, they can contain a considerable amount of detail and it is not always easy to know what is important, not to be sure that a comparison of means will be sufficiently informative for a particular application." The approach we follow in the present paper does lead to predictions which are defined by probability distributions, but does not suffer from the deficiency of having insufficient information for applications.

For definiteness, we consider a large parametric class of models of phenomena in nonlinear continuum mechanics, particularly nonlinear elastostatics, where we expect to have information on molecular structures prevalent in manufacturing processes. This information, paradoxically, pertains to events at spatial scales quite different from those on which the principal continuum-mechanics model is based. We must, therefore, cope with the problem of multiscale validation. We address this problem in the context of general model classes depicting finite deformation of polymeric structures.

In Section 2, following this Introduction, we describe a class of problems in nonlinear elastostatics that provides a model for the discussion of the process of prediction of physical events. There we outline steps in a fabrication process involving polymeric materials. This specific example is one typical of many such processes, the principal feature being that some prior information is known about the structure of the problems at scales smaller than those normally involved in the prediction model. Details on the structure of typical molecular models are given in an appendix to the body of the paper. In Section 3, we review the framework for Bayesian methods for statistical inverse problems arising in calibration and validation processes together with basic ideas of information theory relevant to model validation. In Section 4, we present a model validation process, and propose validation criteria. During the validation process, calibration and validation data are collected from experiments performed at the corresponding scenarios. The validation process leads naturally to the discussion of experiment design in Section 5, where we propose extra validation criteria formulated in terms of information gain, that is, the expected Kullback-Leibler (KL) divergence between probability density functions (pdfs). In Section 6, we come to the practical issue of selecting the validation scenario among

candidate scenarios, in the context of nonlinear elastostatics. The validation scenario selection issue, we assert, can be partially resolved by performing virtual validation experiments (i.e., simulating the validation data) using multiscale models generated with the goal-oriented multiscale algorithms developed by Oden, Bauman, Prudhomme, et al [8, 9, 49, 50, 51]. Virtual experimentation, we argue, provides a basis for the design of actual validation experiments as well as a means for assessing model bias. We also describe how hierarchical model classes and the calculation of model plausibilities can be used in some cases to reduce model bias. Concluding remarks are collected in Section 8.

2. A Model Class of Multiscale Models: Parametric Classes of Models in Nonlinear Elastostatics

The general class of problems we address here involves the use of complex multiscale models for predicting important quantities of interest (QoIs) critical to important decisions, including design decisions.

2.1. Microscale Models of Polymeric Materials

For definiteness, we begin with a focus on macroscale models of the mechanical behavior of solid bodies for which something is known about their microstructure because of prior information on how they were fabricated. One important class of such problems includes the development of models of the mechanical behavior of polymeric components such as those used in some circuit boards, flexible pipes, shock absorbers, lithography etching barriers, etc. The following fabrication process and modeling considerations are common for these types of systems:

1) A (relatively) rigid mold or template is filled with a photocurable solution or mixture consisting of various monomer molecules, cross linkers, initiators, reactants, and voids, each constituent having known initial volume fraction. An energy source, such as ultraviolet light, is applied to the mixture, creating chemical reactions that produce stable polymer chains and thereby a polymeric solid. The process is often based on the assumption that the reaction rate κ obeys an Arrhenius-type law,

$$\kappa = A \exp(-E_a/k_B T)\,, \tag{1}$$

where E_a is the activation energy, the threshold energy at which a reaction, *e.g.* a chemical bonding, takes place, k_B is Boltzmann's constant, T is the (local) absolute temperature, and A is a constant. The most likely reaction between monomers and reactants at a lattice site in a polymer lattice will correspond to the molecule-molecule pairs sharing the largest reaction rates. The probability of a reaction is assumed to be proportional to this reaction rate. This is the basis of the polymerization algorithm where a sequence of reactions leads to polymer molecular chains in a realization of the actual polymer molecular structure. A single realization of the charge-neutral molecular structure of the body (or a close approximation to it) is established at the conclusion of this *polymerization process*. This so-called kinetic Monte Carlo process leading to the polymerization conformation is described in [8] and [9] and is adapted from a method proposed by Burns et al [13]. For a general discussion, see Voter [63]. Based on each such molecular conformation, molecular energy potentials are assigned to each site to establish various bond-types and potentials for defining intermolecular forces on each site, the potentials being selected from standard characterizations of covalent and polar bonds for the final molecular conformation for each realization.

2) The polymerized material is removed from the mold and the molecules quickly move to a state of static equilibrium in which all molecular forces are balanced. This is called the *densification* step. This fabrication process produces geometrically-complex polymeric components, such as conceptualized in Figure 1. If $\mathbf{u}_i$ denotes the displacement of a molecule site i (e.g. a monomer), then the equilibrium configuration of the body is determined by solutions of large nonlinear algebraic systems of the form

$$\frac{\partial V(\boldsymbol{\mu}; \mathbf{u}^n)}{\partial \mathbf{u}_i} - \mathbf{f}_i = 0, \quad i = 1, 2, \ldots, n, \tag{2}$$

where $V(\boldsymbol{\mu}; \mathbf{u}^n)$ is the total potential energy of the molecular system, $\boldsymbol{\mu}$ being a vector of physical paramters of the interaction potentials for the system, $\mathbf{u}^n = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ the list of unknown site displacement vectors of the n molecular sites, and $\mathbf{f}_i$ the applied force at site i. Additional details on the possible forms of $V(\boldsymbol{\mu}; \mathbf{u}^n)$ and examples of the parameters $\boldsymbol{\mu}$ are given in the Appendix A. In general, for the full molecular model of the polymer, n is a very large, often $\mathcal{O}(10^{10})$ or more, so the actual solution of (2) is often difficult and often impossible using contemporary computing capabilities.

3) If the constituents of the mixture are chosen so that the polymer ultimately generated is insensitive to mechanical rate effects (e.g. creep, viscoelastic response), and there are no directional or positional biases in the polymerization process, then the resulting material can be assumed to be elastic, isotropic and homogeneous at the macro scale. The strains resulting from the densification step can be large and significant volumetric changes can occur. All of these considerations suggest that the polymeric structure can be modeled on a macroscale level as an isotropic, homogeneous, compressible, hyperelastic material.

2.2. A Class of Continuum Models

Moving now from the class of molecular models of a coarse-grained characterization of each polymer realization, we consider instead a macro-scale, continuum model of such a hyperelastic body where material coefficients are obtained through some averaging process for each realization. Consider a homogeneous, isotropic, hyperelastic body occupying a reference configuration $\Omega_0 \subset \mathbb{R}^3$, and in static equilibrium under the action of prescribed forces. The motion (or deformed configuration) of the body is determined by the positions, $\mathbf{x}+\mathbf{u}(\mathbf{x})$, of material particles $\mathbf{x}$, $\mathbf{u} = \mathbf{u}(\mathbf{x})$ being the displacement of $\mathbf{x}$. The deformation tensor at $\mathbf{x}$ is

$$\mathbf{F}(\mathbf{x}) = \mathbf{I} + \boldsymbol{\nabla}\mathbf{u}(\mathbf{x}), \tag{3}$$

$\boldsymbol{\nabla}$ being the material gradient operator, and the right Cauchy-Green deformation tensor $\mathbf{C}$ is defined by

$$\mathbf{C} = \mathbf{C}(\mathbf{u}) = \mathbf{F}^T\mathbf{F} = \mathbf{I} + \boldsymbol{\nabla}\mathbf{u} + \boldsymbol{\nabla}\mathbf{u}^T + \boldsymbol{\nabla}\mathbf{u}^T\boldsymbol{\nabla}\mathbf{u} \tag{4}$$

The mechanical response of such a material body is characterized by a stored energy function W of appropriate deformation measures. For a very broad class of such materials, we can postulate a stored energy function of the form,

$$W = \alpha(I_1 - 3) + \beta(I_2 - 3) + \gamma(J-1)^2 - (2\alpha + 4\beta)\ln J\,, \tag{5}$$

where α, β, γ are material constants and I_1, I_2 and J are principal invariants of the $\mathbf{C}$:

$$I_1 = \operatorname{tr}\mathbf{C}\,; \quad I_2 = \operatorname{tr}\mathbf{C}^2 - (\operatorname{tr}\mathbf{C})^2\,; \quad J = \sqrt{|\mathbf{C}|} = \sqrt{\det\mathbf{C}}\,. \tag{6}$$

The form of the stored energy function in (5) is not arbitrarily selected:

i) Being a function of the invariants (6), it is form-invariant under changes in the observer's frame of reference. Thus, it obeys the Principle of Material Frame Indifference for constitutive equations in continuum mechanics.

ii) The first term, $\alpha(I_1 - 3)$ is the well-known Trealor [62] model of incompressible materials, confirmed by arguments based on kinetic theory and statistical mechanics of rubber-like materials (see, *e.g.* [7, 15, 43].

iii) The addition of the second term leads to the famous Mooney-Rivlin energy, $\alpha(I_1 - 3) + \beta(I_2 - 3)$, used to characterize incompressible homogeneous elastomers for over a half century [37, 51, 56].

iv) The last two terms characterize compressibility of the material; the last representing a type of penalty term that represents the property that the energy becomes unbounded ($\to \infty$) as the volume measure $J \to 0$. The coefficient $(2\alpha + 4\beta)$ was suggested by Ciarlet [15].

v) The function W (owing to linearity in I_1 and I_2 and convex behavior in J) is a convex function of the invariants, which themselves are weakly-sequentially continuous functions of the displacement field on appropriate function spaces. The energy is thus polyconvex - and therefore quasiconvex; it is also coercive. Thus, there exist solutions to the non-convex minimization problem of minimizing $\int W \, d\mathbf{x}$ over suitable spaces of displacements $\mathbf{u}$ [15].

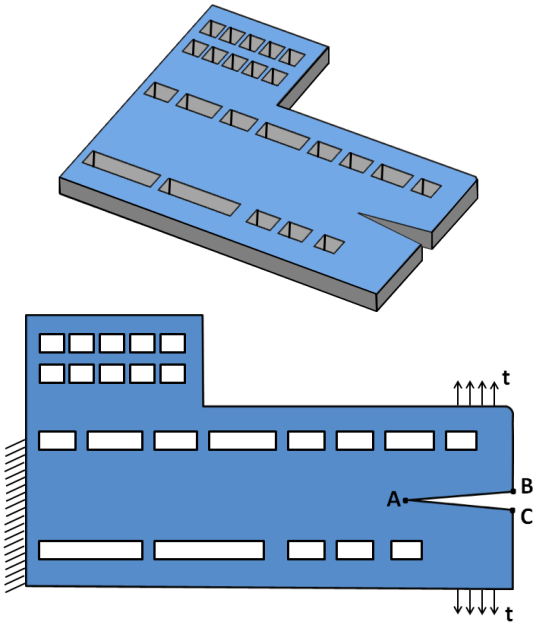


Figure 1: Example of an elastomeric component fabricated through polymerization and in static equilibrium under the action of external tractions.

The local macroscale equilibrium states of the polymer component are characterized by the balance of linear momentum, which is manifested in the static equilibrium conditions charac-

terizing minima of $\int W d\mathbf{x}$:

$$\mathrm{Div}\left(\frac{\partial W}{\partial \mathbf{F}}\right)(\mathbf{x}) = \mathbf{0}\,; \quad \mathbf{x} \in \Omega_0 \subset \mathbb{R}^3\,. \tag{7}$$

Here Div is the material divergence operator, $\mathrm{Div} = \boldsymbol{\nabla}\cdot$. Substitution of (5) and (6) into (7) leads to a system of highly nonlinear partial differential equations (PDEs) in the displacement field $\mathbf{u}$:

$$\begin{aligned} \boldsymbol{\nabla}\cdot\{\ (\mathbf{I}+\boldsymbol{\nabla}\mathbf{u})\cdot\{\ \alpha\mathbf{I}\ + \\ +\ \beta[(\mathrm{tr}\,\mathbf{C}(\mathbf{u})^{-1})\mathbf{I} - \mathbf{C}(\mathbf{u})^{-T}(\mathrm{Cof}\,\mathbf{C}(\mathbf{u}))]\ + \\ +\gamma(\mathrm{Cof}\,\mathbf{C}(\mathbf{u}))\left(1-|\mathbf{C}(\mathbf{u})|^{-1/2}\right)\ + \\ -\ (2\alpha+4\beta)(\mathrm{Cof}\,\mathbf{C}(\mathbf{u}))\ \}\ \} \quad &= \quad 0\,. \end{aligned} \tag{8}$$

These notations and models are standard in nonlinear elasticity and continuum mechanics; see e.g. [7, 15, 23] or the summary account in [43]. The system (8) describes an infinite number of models of the deformation of a hyperelastic body, infinite because each realization of the parameters

$$\boldsymbol{\theta} = (\alpha, \beta, \gamma)\,,$$

for each polymerization realization, corresponds to a different model. Alternatively, the parameters $\boldsymbol{\theta}$ can be regarded as random variables and (8) is then a system of stochastic PDEs with solutions themselves random fields.

To (8), we must add information on the solution domain Ω_0, boundary conditions on $\partial\Omega_0$, and other data, such as possible source terms on the right-hand side of (5). These define the physical scenario s under which the model equations are to be applied. Since the problem domain might depend upon the scenario in which the model is implemented, the notations $\Omega_0(s)$ and $\partial\Omega_0(s)$ are meaningful. We could, of course, choose to represent the discretized initial and boundary data as unknown parameters and include them in the list of parameters $\boldsymbol{\theta}$, but for clarity we will view the scenarios as known and deterministic.

Upon solving (8) for a given s and $\boldsymbol{\theta}$, one computes the solution

$$\mathbf{u} = \mathbf{u}(\boldsymbol{\theta}, s; \mathbf{x}), \quad \mathbf{x} \in \overline{\Omega_0}(s), \tag{9}$$

which can be used to compute modeled observables for various scenarios, including quantities of interest for a prediction scenario s_{p} to be defined later.

2.3. An Example: Polymer Etch Barrier in Imprint Lithography

An interesting example of the fabrication process and resulting structures occurs in the manufacturing of etch barriers for imprinting semiconductor components in Step and Flash Imprint Lithography (SFIL; see [9, 13]). Figure 2 presents the semiconductor components produced by manufacturing the etching barriers which are polymers of dimension of 30-70 nm. A molecular model of the block structure is shown.

3. Review of Background Concepts Related to Forward and Inverse Problems

The model described so far is one example of a general model of the form,

$$A(\boldsymbol{\theta}, s; u(\boldsymbol{\theta}, s)) = 0, \tag{10}$$

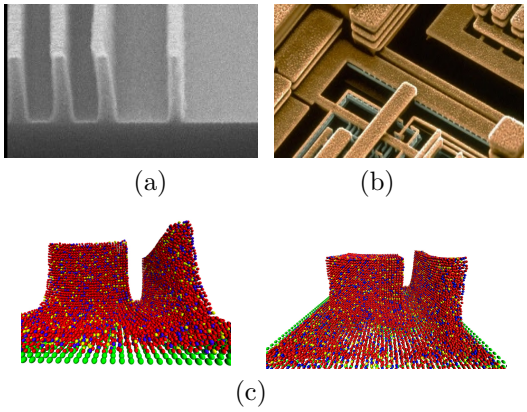


Figure 2: (a) Nanoscale polymeric etch-barrier structures for imprint-lithography semiconductors manufacturing process; (b) semiconductor components; and (c) multiblock molecular model of polymer densification representing etch barrier structures ([9, 13]).

where A denotes an abstract operator such as those in (8), s denotes the scenario (which might also incorporate locations of points or domains where measurements are taken, measurement time window, measurement frequency etc.), and $\boldsymbol{\theta}$ denotes the model parameters.

By a *scenario*, we mean the specific environment in which the broad mathematical model is to be implemented, generally describing the domain of the solutions, the boundary involved in boundary conditions, etc. The prediction scenario s_p corresponds to the full system model used to predict the quantities of interest. It may consist of several subsystems depending on the characteristics of the full model. The simplest scenarios are the calibrations scenarios s_c designed to provide updated information on model parameters (updated from priors). In multiphysics models, different calibration scenarios may be called on to address different physical behaviors (*e.g.*, in a model of fluid-structure interaction, calibration tests may be done to determine parameters defining the model of the structural material while others are used to calibrate the fluid model, and still others to calibrate combined fluid-structure models).The validation scenarios s_v are encountered in subsystem models designed to challenge basic assumptions on which the full system model is based, as well as to update parameters delivered by the calibration tests that can be better informed (*i.e.* that exhibit an information gain) over those obtained through calibrations. The validation scenarios need not be a subset of s_p. They can correspond to independent domains over which the mathematical model is implemented or structured to deliver information on the utility of the full model for predicting the QoIs. The design of validation experiments, and implicitly the validation scenarios, is a major challenge in predictive science; and is an issue taken up in this paper. As a general rule, the scenarios are assumed to involve no uncertainties, but this assumption is easily relaxed. It is generally the case, although exceptions can be cited, that with increasing the sophistication of the validation scenarios (equivalently, increasing the complexity of the subsystem models), there is a decrease in observational data available to inform the model. More specifically, validation experiments may re-tune parameters to account for behavior not captured by calibration tests.

Obviously, model (10) can be applied to many different scenarios s. We define

$$\mathcal{S} = \left\{ \begin{array}{l} \text{the set of all possible scenarios } s \text{ to which} \\ \text{the model can be applied.} \end{array} \right. \tag{11}$$

The random vector of model parameters $\boldsymbol{\theta}$ is regarded as a member of a manifold $V_{\boldsymbol{\theta}}$. We assume that (10) is well-posed and a solution $u(\boldsymbol{\theta}, s)$ exists for every $\boldsymbol{\theta} \in V_{\boldsymbol{\theta}}$ and scenario $s \in \mathcal{S}$. Note that this assumption is satisfied by (8) for $\alpha, \beta, \gamma > 0$ as explained in comment (v) in subsection 2.2. As $\boldsymbol{\theta}$ runs over $V_{\boldsymbol{\theta}}$, (10) characterizes a parametric class of models.

We note that the scenario s can also be treated as a random entity. We note again that, even though the governing operators defining $A(\cdot, \cdot; \cdot)$ may be deterministic, the coefficients may depend on random parameters $\boldsymbol{\theta}$ and random scenario s, so the solution $u(\boldsymbol{\theta}, s)$ is a random field.

3.1. Statistical Inverse Problems and Experimental Data $\mathbf{d}$

Key steps in our approach to model validation are the construction and solution of statistical inverse problems, which are characterized by the Bayesian updates of pdfs π of model parameters. Initially, our knowledge of the parameters is characterized by the prior pdf $\pi(\boldsymbol{\theta})$. The fundamental Bayesian relationship is:

$$\pi(\boldsymbol{\theta}|\mathbf{d}, s) = \frac{\pi(\mathbf{d}|\boldsymbol{\theta}, s)\pi(\boldsymbol{\theta}|s)}{\pi(\mathbf{d}|s)}. \tag{12}$$

Here $\pi(\boldsymbol{\theta}|\mathbf{d}, s)$ is the posterior pdf, the conditional probability on model parameters $\boldsymbol{\theta}$, given scenario s and experimental data $\mathbf{d}$; $\pi(\mathbf{d}|\boldsymbol{\theta}, s)$ is the likelihood pdf, the conditional pdf on $\mathbf{d}$ given $\boldsymbol{\theta}$, generally based on the model employed to relate parameters to observations; $\pi(\boldsymbol{\theta}|s)$ is the prior pdf on $\boldsymbol{\theta}$; and

$$\pi(\mathbf{d}|s) = \int \pi(\mathbf{d}|\boldsymbol{\theta}, s)\pi(\boldsymbol{\theta}|s)\ \mathrm{d}\boldsymbol{\theta},$$

since $\int \pi(\boldsymbol{\theta}|\mathbf{d}, s)\ \mathrm{d}\boldsymbol{\theta} = 1$. The setting is classical: beginning with prior information on the parameters $\boldsymbol{\theta}$, their uncertainty represented by the prior pdf $\pi(\boldsymbol{\theta}|s)$, perform physical experiments (or make observations of physical events) to produce $\mathbf{d}$ with uncertainties depicted in $\pi(\mathbf{d}|s)$. Then compute an updated characterization of $\boldsymbol{\theta}$ in the pdf $\pi(\boldsymbol{\theta}|\mathbf{d}, s)$. The key is the likelihood, which connects $\boldsymbol{\theta}$ and $\mathbf{d}$ through the map provided by the mathematical model, the model of uncertainties in the observational data, and the model of inadequacies in the model itself.

The theoretical model (10) is thus appended with the calculations of theoretical predictions $\mathbf{y}_s$ of the observations $\mathbf{d}$ for any particular scenario s,

$$\left.\begin{array}{rcl} A(\boldsymbol{\theta}, s; u(\boldsymbol{\theta}, s)) & = & 0, \\ \mathbf{y}_s & = & \mathbf{y}(\boldsymbol{\theta}, s, u(\boldsymbol{\theta}, s)), \end{array}\right\} \tag{13}$$

where $\mathbf{y}(\boldsymbol{\theta}, s, u(\boldsymbol{\theta}, s))$ is the vector of values computed by the model, which are to be compared to the observable data $\mathbf{d}$. Such comparison involves extra assumptions, which are encapsulated into the likelihood pdf $\pi(\mathbf{d}|\boldsymbol{\theta}, s)$ [14, 48]. For instance, is the discrepancy between $\mathbf{y}$ and $\mathbf{d}$ additive? Or maybe multiplicative? What is the covariance structure of such discrepancy? Is a simple diagonal covariance matrix a viable representation? The modeling of the discrepancy may also contribute additional unknown parameters to the process. The subject of model discrepancy is discussed further in Sections 4 and 7.

For the case that s is the prediction scenario, $s = s_{\mathrm{p}}$, and for parameters $\boldsymbol{\theta}$, the computed solution $u = u(\boldsymbol{\theta}, s_{\mathrm{p}}; \mathbf{x})$ is used to compute a real-valued QoI, q, that defines the overall goal of the modeling exercise. The model (10) is thus augmented as follows:

$$\left.\begin{array}{rcl} A(\boldsymbol{\theta}, s_{\mathrm{p}}; u(\boldsymbol{\theta}, s_{\mathrm{p}})) & = & 0, \\ q & = & q(\boldsymbol{\theta}, s_{\mathrm{p}}, u(\boldsymbol{\theta}, s_{\mathrm{p}})). \end{array}\right\} \tag{14}$$

For the hypothetical elastomer component shown in Figure 1, examples of QoIs are the maximum principal value of the Cauchy stress at the point A or the displacement separation of points B and C. The QoIs are not observables; they are extrapolations of what is known about the model and the experimental data to future events, not measured in experimental set ups. Once a quantity is measured it becomes part of the experimental calibration or validation data. We will refer to problem (14) as the forward problem.

3.2. Information Related Concepts

For future reference, we record useful measures of information and uncertainty. See [19] for a complete treatment. Let X, Y be random variables with pdfs $\pi(x)$ and $\pi(y)$, and let $\pi(x, y)$ denote the joint pdf of X and Y.

3.2.1. Information Entropy, Conditional Entropy, and Relative Entropy

Arguments of Jaynes [27] building on the classical work of Shannon [59], show that the only reasonable measure of the amount of uncertainty embodied in a random variable that satisfies certain mild consistency conditions is its information entropy,

$$H(X) = \text{“}H(\pi)\text{”} = -\int \pi(x) \log \frac{\pi(x)}{m}\, dx, \tag{15}$$

m being an invariant measure.

The conditional entropy of X, given Y, is defined as

$$\begin{aligned} H(X|Y) &= -\int\int \pi(x, y) \log \pi(x|y)\, dx\, dy \\ &= H(X, Y) - H(Y). \end{aligned} \tag{16}$$

The relative entropy between pdfs $\pi_2(x)$ and $\pi_1(x)$, also called the Kullback-Leibler (KL) divergence between the two pdfs, is defined as

$$D_{KL}(\pi_2||\pi_1) = \int \pi_2(x) \log \frac{\pi_2(x)}{\pi_1(x)}\, dx \; = -H(\pi_2) + H(\pi_2, \pi_1). \tag{17}$$

The KL-divergence is thus the sum of the negative informational entropy of π_2 and the so-called cross entropy $H(\pi_2, \pi_1) = -\int \pi_2(x) \log \pi_1(x)\, dx$. If, for example, $\pi(\boldsymbol{\theta}|\mathbf{d})$ is the posterior update of a prior $\pi(\boldsymbol{\theta})$, given data $\mathbf{d}$, the expected value of the KL-divergence between the posterior and prior integrated over the data, is the information gain; $I(\pi(\boldsymbol{\theta}|\mathbf{d}), \pi(\boldsymbol{\theta})) = \int\int D_{KL}(\pi(\boldsymbol{\theta}|\mathbf{d})\|\pi(\boldsymbol{\theta}))\pi(\mathbf{d})\, d\boldsymbol{\theta} d\mathbf{d} = \mathbb{E}_{\mathbf{d}}(D_{KL}(\pi(\boldsymbol{\theta}|\mathbf{d})\|\pi(\boldsymbol{\theta})))$.

3.2.2. Mutual Information and Conditional Mutual Information

One measure of the relative information contained in X and Y is the Shannon mutual information [19],

$$I(X; Y) = \int \int \pi(x, y) \log \frac{\pi(x, y)}{\pi(x)\pi(y)}\, dx\, dy, \tag{18}$$

which can be easily shown to satisfy

$$\begin{aligned} I(X; Y) &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \\ &= H(X) + H(Y) - H(X, Y). \end{aligned} \tag{19}$$

The mutual information is a measure of independency between X and Y, since it is zero if the random variables are independent. It can also be seen as a measure of the amount of information that one random variable contains about another random variable [19, p. 19].

Let Z be a random variable. The conditional mutual information of X and Y, given Z, is defined as

$$I(X;Y|Z) = H(X|Z) - H(X|Y,Z). \tag{20}$$

Henceforth we do not distinguish between an information measure of a random variable X and its pdf, writing, for example, $H(\pi)$ for $H(X)$, a convention used throughout [27].

4. A Model Validation Process

As explained in [5, 45, 46], model validation is a process designed to produce information leading to confidence, or lack of confidence, in the ability of the model to predict QoIs. As noted in the Introduction, the process can only establish if the model is not invalid under certain conditions relative to assigned tolerances of the "degree of validity", as well as to given experimental data. We lay out here a process inspired by [6] with some generalizations. The process involves the following scenarios:

- the calibration scenarios $s_{\rm c}$, which generally corresponds to simple laboratory component experiments that yield the calibration data $\mathbf{d}_{\rm c} = \mathbf{d}(s_{\rm c})$, and
- the validation scenarios $s_{\rm v}$, which correspond to more complex sub-system-level experiments that yield the validation data $\mathbf{d}_{\rm v} = \mathbf{d}(s_{\rm v})$.

In general, there may be several calibration experiments, particularly in multiphysics modeling, wherein different physics attributes and their key parameters are updated in different calibration experiments. Calibration, by definition, involves determining model parameters to best fit experimental data. Validation experiments are, in principle, designed to challenge the model, to test the validity of assumptions made in developing the model and how they affect the prediction of the QoIs. For clarity, it is sufficient here to consider a single calibration scenario and experiment, but we must face the problem of considering several possible validation experiments: $s_{\rm v1}, s_{\rm v2}, \ldots, s_{\rm vK}$. Indeed, a *central problem in model validation is the selection of the validation scenarios that provide the most information on the capability of the model to predict the QoI with acceptable accuracy (its "validity"), subject to some constraints*, such as cost, feasibility, etc. The prediction scenario $s_{\rm p}$ prevails when carrying out the actual goal of the simulation: the calculation of the target output, the QoI, Q.

As indicated in (13), for any scenario s, the goal is not simply to solve the forward problem for $u(\boldsymbol{\theta}, s)$, but is to compute predictions $\mathbf{y}$ to be compared with experimental measurements, the observations $\mathbf{d}$. For scenario s, we arrange theoretical outputs $\mathbf{y}_s(\boldsymbol{\theta}) = \mathbf{y}(u(\boldsymbol{\theta}, s))$ in a vector to be compared with the observations $\mathbf{d} = \mathbf{d}_s$ corresponding to s. In general, the measurements $\mathbf{d}_s$ agree neither with $\mathbf{y}$ nor with the actual physical realities. Let $\boldsymbol{\Lambda}_s$ denote the true physically real values of the outputs for s (the unobservable realities). Then we have

$$\boldsymbol{\Lambda}_s = \mathbf{d}_s - \boldsymbol{\varepsilon}_s, \tag{21}$$

where $\boldsymbol{\varepsilon}_s$ is the experimental error or noise. According to [10, 32, 33], the difference between the real values of the physical entities and those delivered by the model is defined as $\mathbf{b}_s$, the *model bias* (or *model inadequacy* or *model discrepancy* or *model form error*),

$$\mathbf{b}_s = \boldsymbol{\Lambda}_s - \mathbf{y}_s. \tag{22}$$

Thus,

$$\mathbf{d}_s = \mathbf{y}_s(\boldsymbol{\theta}) + \mathbf{b}_s + \boldsymbol{\varepsilon}_s. \tag{23}$$

Clearly, to determine the distribution of $\mathbf{d}_s$ given $\boldsymbol{\theta}_s$, one must know or have a model of the total discrepancy, $\mathbf{b}_s + \boldsymbol{\varepsilon}_s$; we refer to models of this sum as *uncertainty models*. We return to this concept later, in Section 6.

At this point, the symbolic notion of a *prediction pyramid* discussed in [17, 44, 45, 46] is almost unavoidable. As shown in Figure 3, the lowest level of the prediction process begins at the base of the pyramid with calibration scenarios s_c and calibration data $\mathbf{d}_\mathrm{c}$. The next level is characterized by validation scenarios s_v and validation data $\mathbf{d}_\mathrm{v}$. Finally, the computationally most demanding scenario is manifested in the prediction scenario s_p, in which the target QoI, Q, is calculated. As we move up the pyramid, generally the amount of data decreases as the complexity of the model and experiments increases. To complete the metaphor, the hidden face of the pyramid can be the locations of the unobservables: the realities $\boldsymbol{\Lambda}_s$.

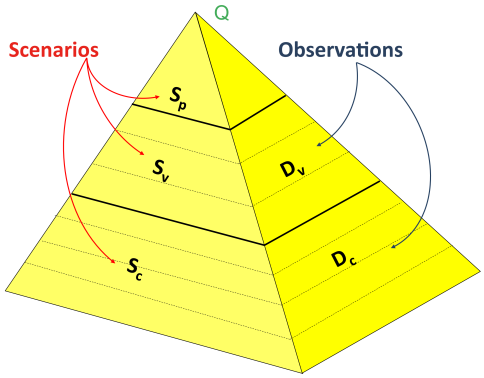


Figure 3: The prediction pyramid (Cf. [17, 44, 45, 46]). For each scenario s, the choices $\mathbf{d}$ of observations are made. The unseen face of the pyramid may be regarded as analogous to the unobservables - the physical realities, $\boldsymbol{\Lambda}_s$, of the outputs.

We remark that such segregating of validation processes into phases of increasing complexity is not at all new in the literature. In the 1998 Guide [1], for example, a hierarchy of validated phases is given, beginning with low level unit problems, then benchmark cases, then subsystem cases, and finally the complete system. The unit problems roughly correspond to the calibration scenario, the benchmark cases, to low-level validation scenarios designed to test key assumptions; the subsystem cases to high-level validation scenarios designed to depict features that are influenced by the QoIs. The complete system is the full prediction scenario.

4.1. The Calibration Phase

Given the prior information $\pi(\boldsymbol{\theta})$ on the model parameters, and the observational data $\mathbf{d}_\mathrm{c}$ obtained in calibration experiments at scenario s_c, we update the joint pdf of the model parameters via

$$\pi(\boldsymbol{\theta}|\mathbf{d}_\mathrm{c}) = \frac{\pi(\mathbf{d}_\mathrm{c}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{d}_\mathrm{c})}. \tag{24}$$

Let us denote by $\boldsymbol{\Theta}_\mathrm{c}$ the random variable characterized by the pdf (24). Any possible realization $\boldsymbol{\theta}$ of such a random variable leads, via the forward problem (14), to a corresponding QoI, q,

We define the random variable $Q_{\rm c}$ as the calibration QoI characterized by the pdf

$$\pi(q|\mathbf{d}_{\rm c}) = \int \pi(q|\boldsymbol{\theta}, \mathbf{d}_{\rm c})\pi(\boldsymbol{\theta}|\mathbf{d}_{\rm c})\ \mathrm{d}\boldsymbol{\theta}. \tag{25}$$

We employ sampling algorithms such as MCMC [48, 14, 54, 55] in order to generate samples of $\boldsymbol{\Theta}_{\rm c}$, and the computation of $\pi(q|\mathbf{d}_{\rm c})$ can be done using any of several approaches, Monte Carlo being the classical one, but others, such as polynomial chaos, stochastic collocation etc., could be used (see, e.g. [34]).

The calibration phase provides the important opportunity to test the basic hypotheses on which the model is based. In the case of model (8), one would wish to test the extent to which the assumptions of a homogeneous, isotropic, and hyperelastic material stood up to calibration experiments. Some of these hypotheses can be further tested in the validation phase, explained next.

The information entropy (15) can be used to quantify the uncertainty in $Q_{\rm c}$:

$$H(Q_{\rm c}) = -\int \pi(q|\mathbf{d}_{\rm c}) \log \frac{\pi(q|\mathbf{d}_{\rm c})}{m}\ dq. \tag{26}$$

4.2. The Validation Phase

In addition to running the calibration phase, which consists of calibration experiments performed at calibration scenarios, a Bayesian inference, and the prediction of the random variable $Q_{\rm c}$, we run a validation phase in a similar fashion. Firstly, we perform one (or more) validation experiment(s) at validation scenario(s) $s_{\rm v}$ (or $s_{\rm v1}, s_{\rm v2}, \ldots$), collecting data $\mathbf{d}_{\rm v}$ (or $\mathbf{d}_{\rm v1}, \mathbf{d}_{\rm v2}, \ldots$). Secondly, we perform an extra Bayesian update on the model parameters, using as prior the calibrated parameter joint pdf (24), that is,

$$\pi(\boldsymbol{\theta}|\mathbf{d}_{\rm v}, \mathbf{d}_{\rm c}) = \frac{\pi(\mathbf{d}_{\rm v}|\boldsymbol{\theta}, \mathbf{d}_{\rm c})\pi(\boldsymbol{\theta}|\mathbf{d}_{\rm c})}{\pi(\mathbf{d}_{\rm v}|\mathbf{d}_{\rm c})}. \tag{27}$$

Let us denote by $\boldsymbol{\Theta}_{\rm v}$ the random variable characterized by the pdf (27). Thirdly, and last, any possible realization $\boldsymbol{\theta}$ of such random variable gives, via (10), to a corresponding Q quantity, whose pdf is indicated by $\pi(q|\boldsymbol{\theta}, \mathbf{d}_{\rm v}, \mathbf{d}_{\rm c})$, and we denote by $Q_{\rm v}$ the random variable characterized by the pdf

$$\pi(q|\mathbf{d}_{\rm v}, \mathbf{d}_{\rm c}) = \int \pi(q|\boldsymbol{\theta}, \mathbf{d}_{\rm v}, \mathbf{d}_{\rm c})\pi(\boldsymbol{\theta}|\mathbf{d}_{\rm v}, \mathbf{d}_{\rm c})\ \mathrm{d}\boldsymbol{\theta}. \tag{28}$$

The quantity of uncertainty in $Q_{\rm v}$ is

$$H(Q_{\rm v}) = -\int \pi(q|\mathbf{d}_{\rm v}, \mathbf{d}_{\rm c}) \log \frac{\pi(q|\mathbf{d}_{\rm v}, \mathbf{d}_{\rm c})}{m}\ dq \tag{29}$$

4.3. Comparing the QoI Predictions

At this point there are four issues that come to the forefront:

1. As we progress up the prediction pyramid, with successive validation experiments and decreasing observational data, have we gained information relevant to the predictability of the QoIs?

2. If so, what is the quantity of uncertainty in our prediction of the QoI?

3. Is the estimated quantity of uncertainty acceptable: *i.e.* is the model acceptable for its intended use?

4. Is our best prediction of the QoI sufficiently accurate to declare that the model is not invalid?

The first question has to do with the question of "over fitting" data - using data that are statistically the same on nearly the same to infer new information on model parameters through validation experiments. This becomes an issue of experimental design: design a sequence of validation experiments so information is gained on the parameters relevant to the QoIs (see, *e.g.* [16, 39, 58]. We discuss this more fully in the next section.

Given a sequence of K validation experiments, denote the posterior pdf at the i-th validation level by

$$\pi(\boldsymbol{\theta}|\mathbf{d}_{\mathrm{v}i}, \mathbf{d}_{\mathrm{v}i-1}, \dots \mathbf{d}_{\mathrm{c}}), \tag{30}$$

and let Q_i denote the predicted QoI obtained from the solution of the forward problem using the parameter posterior. The pdf of Q_i is

$$\pi_i(q) = \int \pi(q|\boldsymbol{\theta}, \mathbf{d}_{\mathrm{v}i}, \mathbf{d}_{\mathrm{v}i-1}, \dots, \mathbf{d}_{\mathrm{c}})\ \pi(\boldsymbol{\theta}|\mathbf{d}_{\mathrm{v}i}, \mathbf{d}_{\mathrm{v}i-1}, \dots, \mathbf{d}_{\mathrm{c}})\ d\boldsymbol{\theta}, \quad i = 1, 2, \dots, K, \tag{31}$$

and its uncertainty content is

$$H(Q_i) = -\int \pi_i(q) \log \frac{\pi_i(q)}{m}\ dq, \quad i = 1, 2, \dots, K. \tag{32}$$

This addresses issue 2. Since the posterior of the i-th validation level is used as the prior at the $(i+1)$-th level, we have

$$H(Q_K) \leq H(Q_{K-1}) \leq \dots \leq H(Q_1). \tag{33}$$

At the conclusion of the validation process, we compute

$$\pi(\boldsymbol{\theta}|\mathbf{d}_K, \dots, \mathbf{d}_{\mathrm{c}}) = \frac{\pi(\mathbf{d}_K|\boldsymbol{\theta}, \mathbf{d}_{K-1}, \dots, \mathbf{d}_{\mathrm{c}})\ \pi(\boldsymbol{\theta}|\mathbf{d}_{K-1}, \dots, \mathbf{d}_{\mathrm{c}})}{\pi(\mathbf{d}_K|\mathbf{d}_{K-1} \dots, \mathbf{d}_{\mathrm{c}})} \tag{34}$$

and solve the forward problem in the prediction scenario for $Q = Q_{\mathrm{p}}$, the prediction QoI, and we compute its measure of uncertainty $H(Q_{\mathrm{p}})$.

Returning to issue 3, the decision as to the acceptability of the uncertainty in the predicted QoI must be determined by the decision maker. Given a tolerance η_0 for uncertainty, the decision maker will deem the prediction acceptable for its intended use whenever $H(Q_{\mathrm{p}}) \leq \eta_0$. The determination of η_0 requires other considerations outside the validation process (using, *e.g.*, loss functions, decision theory, or heuristics).

Returning now to the original model based on the use of the calibrated parameters and the prediction scenario s_{p}, the original question, now somewhat lost in the validation process, was whether or not the original model with calibrated parameters has a degree of validity sufficient for the prediction of the QoI with a level of uncertainty acceptable for its intended use? We can only compare its predictability with that of the products of the validation process which employed all of the data available to the analyst. In many respects, while interesting and relevant to the original goal of validating a given model, this question is often of secondary importance as the outcome of a successful validation process is the prediction of QoI distributions with acceptable levels of uncertainty. If the validation process involved well-designed experiments and

adequately informed the model on the QoI, then the final K-th level stands as the last step in updating the parameters. Let us assume that our estimate of $H(Q_{\rm p})$ meets our preset tolerance for validity. Then the question of validity of the original model reduces to a comparison of the calibrated and validated models. The KL divergence between the two QoI pdfs is

$$D_{\rm KL}[\pi(q|\mathbf{d}_{\rm v},\mathbf{d}_{\rm c})||\pi(q|\mathbf{d}_{\rm c})] = \\ = \int \pi(q|\mathbf{d}_{\rm v},\mathbf{d}_{\rm c}) \log \frac{\pi(q|\mathbf{d}_{\rm v},\mathbf{d}_{\rm c})}{\pi(q|\mathbf{d}_{\rm c})}\, {\rm d}q. \tag{35}$$

So, given the deterministic vectors (realizations of random vectors) $\mathbf{d}_{\rm v}$ and $\mathbf{d}_{\rm c}$, we choose a tolerance $\eta_{\rm tol}$ and declare:

$$\left.\begin{array}{rcl} \text{The calibrated model is } \underline{\textbf{invalid}} \text{ if } D_{\rm KL} \text{ in (35)} & > & \eta_{\rm tol}, \\ \text{The calibrated model is } \underline{\textbf{not invalid}} \text{ if } D_{\rm KL} \text{ in (35)} & \leq & \eta_{\rm tol}. \end{array}\right\} \tag{36}$$

It is important to understand the reasoning behind the validation test (36). We are given a model with parameters calibrated to produce the posterior pdf $\pi(\boldsymbol{\theta}|\mathbf{d}_{\rm c})$. The question is: is this a valid model for computing the QoI for the prediction scenario $s_{\rm p}$? The validation process is viewed as a separate set of considerations designed to test the validity of the calibrated model – not to predict Q. On the other hand, $\pi(q|\mathbf{d}_{\rm v},\mathbf{d}_{\rm c})$ is expected to be a better characterization of Q and the uncertainty in Q, since it is conditioned on the conjunction of the calibration data $\mathbf{d}_{\rm c}$ and validation data $\mathbf{d}_{\rm v}$. Thus, one could, and often should, use the validation prediction $\pi(q|\mathbf{d}_{\rm v},\mathbf{d}_{\rm c})$ and the characterization of Q and its uncertainty. But then this would have nothing to do with the question of whether the original calibrated model is valid, nor would it confirm if any of the validation steps or their ultimate prediction was based on valid models.

It is also important to note that (36) can be interpreted as a comparison of a prediction of a QoI delivered by the calibrated model in the prediction scenario defined by the pdf $\pi(q|\mathbf{d}_c)$, with the pdf $\pi(q|\mathbf{d}_v,\mathbf{d}_c)$, delivered by the model, also in the prediction scenario, but with new parameter distributions that have been informed (updated) using the highest-level validation data. This is a generalization of the primitive validation experiments in which deterministic measurements are compared with model predictions, except that in the present framework the quantities being compared are pdfs, all having been conditioned to the available data, and with data having uncertainties. Note also that the D_{KL} of (35) measures the relative entropy of the prediction pdf and that of the validation process (recall (17)).

For a given fixed tolerance, η_{tol}, the criterion (36) may lead to the rejection of the calibrated model as it compares the outcome of such model with that of a validation model that used a richer data set. An alternative would be to consider the convergence of a sequence of predicted QoIs, $\pi(Q_k)$, $k = 0, 1, \ldots, K$, with $\pi(Q_0) = \pi(Q_c)$, and to judge the quality of successive predictions by computing

$$\gamma_k = D_{KL}(\pi(Q_k)||\pi(Q_{k-1})), \quad k \geqslant 1. \tag{37}$$

If γ_K is small enough ($\gamma_K \leq \gamma_{tol}$), then the model can be declared not invalid. This presumes that the validation data $\mathbf{d}_{vK}$ is selected to properly influence the QoI, an issue of experimental design of the validation experiments. These ideas are offered as an answer to question 4.

5. Design of Experiments

At this point, it is natural to ask the question:

$$\left.\begin{array}{l}\text{How much new information on the predictions have}\\ \text{the (sampled) validation data } \mathbf{d}_{\mathrm{v}} \text{ furnished beyond}\\ \text{that supplied by the calibration data } \mathbf{d}_{\mathrm{c}}?\end{array}\right\} \tag{38}$$

Such a question naturally relates itself to the KL divergence (35) between the two QoI pdfs (24) and (27). So, a natural follow-on question is:

$$\left.\begin{array}{l}\text{What is the validation scenario } s_{\mathrm{v}} \text{ that produces}\\ \text{the (random) validation data } \mathbf{d}_{\mathrm{v}} \text{ that maximize}\\ \text{the } \textit{expected information gain} \text{ between the two}\\ \text{QoI pdfs?,}\end{array}\right\} \tag{39}$$

where the expected information gain is defined as

$$\begin{aligned}\mathbb{E}[D_{\mathrm{KL}}] &= \int_{\mathbf{D}_{\mathrm{v}}}\left[\int_{Q}\pi(q|\mathbf{d}_{\mathrm{v}},\mathbf{d}_{\mathrm{c}})\cdot\right.\\ &\quad\left.\cdot\log\frac{\pi(q|\mathbf{d}_{\mathrm{v}},\mathbf{d}_{\mathrm{c}})}{\pi(q|\mathbf{d}_{\mathrm{c}})}\,\mathrm{d}q\right]\pi(\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})\,\mathrm{d}\mathbf{d}_{\mathrm{v}}.\end{aligned} \tag{40}$$

We use the expected information gain, instead of simply information gain, because, as highlighted before, $\mathbf{d}_{\mathrm{v}} = \mathbf{d}(s_{\mathrm{v}})$ is a random vector.

Yet another related question is

$$\left.\begin{array}{l}\text{Given the (sampled) calibration data } \mathbf{d}_{\mathrm{c}}, \text{ how much}\\ \text{can the (random) validation data } \mathbf{d}_{\mathrm{v}} \text{ affect } Q?\end{array}\right\} \tag{41}$$

Such a question involves the concept of conditional mutual information, presented in Subsection 3.2.2. In the situation under study, we have:

$$\begin{aligned}I(Q;\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}}) &= \int_{\mathbf{D}_{\mathrm{v}}}\int_{Q}\pi(q,\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})\cdot\\ &\quad\cdot\log\frac{\pi(q,\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})}{\pi(q|\mathbf{d}_{\mathrm{c}})\pi(\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})}\,\mathrm{d}q\,\mathrm{d}\mathbf{d}_{\mathrm{v}}.\end{aligned} \tag{42}$$

The following theorem is a straightforward consequence of the definitions given thus far.

Theorem 1. *The conditional mutual information* (42) *is the expected value of the KL divergence* (35) *under the pdf* $\pi(\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})$*:*

$$I(Q;\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}}) = \mathbb{E}\{D_{KL}[\pi(q|\mathbf{d}_{\mathrm{v}},\mathbf{d}_{\mathrm{c}})||\pi(q|\mathbf{d}_{\mathrm{c}})]\}. \tag{43}$$

Proof. We have

$$\begin{aligned}\mathbb{E}[D_{\mathrm{KL}}] &= \int_{\mathbf{D}_{\mathrm{v}}}\left[\int_{Q}\pi(q|\mathbf{d}_{\mathrm{v}},\mathbf{d}_{\mathrm{c}})\cdot\right.\\ &\quad\left.\cdot\log\frac{\pi(q|\mathbf{d}_{\mathrm{v}},\mathbf{d}_{\mathrm{c}})}{\pi(q|\mathbf{d}_{\mathrm{c}})}\,\mathrm{d}q\right]\pi(\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})\,\mathrm{d}\mathbf{d}_{\mathrm{v}}.\end{aligned}$$

But

$$\begin{aligned}\pi(q|\mathbf{d}_{\mathrm{v}},\mathbf{d}_{\mathrm{c}}) &= \frac{\pi(q,\mathbf{d}_{\mathrm{v}},\mathbf{d}_{\mathrm{c}})}{\pi(\mathbf{d}_{\mathrm{v}},\mathbf{d}_{\mathrm{c}})}\\ &= \frac{\pi(q,\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})\pi(\mathbf{d}_{\mathrm{c}})}{\pi(\mathbf{d}_{\mathrm{v}},\mathbf{d}_{\mathrm{c}})}\\ &= \frac{\pi(q,\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})}{\pi(\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})}.\end{aligned}$$

Thus,

$$\begin{aligned}\mathbb{E}[D_{\mathrm{KL}}] &= \int_{\mathbf{D}_{\mathrm{v}}}\left[\int_{Q}\frac{\pi(q,\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})}{\pi(\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})}\cdot\right.\\ &\quad\left.\cdot\log\frac{\pi(q,\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})}{\pi(\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})\pi(q|\mathbf{d}_{\mathrm{c}})}\,\mathrm{d}q\right]\pi(\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})\,\mathrm{d}\mathbf{d}_{\mathrm{v}}\\ &= \int_{\mathbf{D}_{\mathrm{v}}}\int_{Q}\pi(q,\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})\cdot\\ &\quad\cdot\log\frac{\pi(q,\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})}{\pi(\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}})\pi(q|\mathbf{d}_{\mathrm{c}})}\,\mathrm{d}q\,\mathrm{d}\mathbf{d}_{\mathrm{v}}\\ &= I(Q;\mathbf{d}_{\mathrm{v}}|\mathbf{d}_{\mathrm{c}}).\end{aligned}$$

□

Therefore, the scenario s_{v}^* that maximizes the expected information gain from Q_{c} to Q_{v} is the same scenario s_{v}^* that maximizes the conditional mutual information between Q and $\mathbf{d}_{\mathrm{v}}$, conditioned on $\mathbf{d}_{\mathrm{c}}$. The best scenario s_{v}^* can thus be seen as the solution of the optimization problem:

$$\max_{s_{\mathrm{v}}\in\mathcal{S}} \quad I(Q;\mathbf{d}_{\mathrm{v}}(s_{\mathrm{v}})|\mathbf{d}_{\mathrm{c}}), \tag{44}$$

subject, of course, to feasibility constraints (e.g. cost) and the assumption that $I(\cdot)$ is bounded.

In practice the scenario set $\mathcal{S}$ is parametrized via $n_\xi \geqslant 1$ bounded design variables $\xi_1,\ldots,\xi_{n_\xi}$. Each design variable can be either discrete or continuous. If at least one is continuous, then the number of possible scenarios is obviously infinite. For example, one can fix the physical scenario and try to optimize the experimental scenario, e.g. find the positions of five available probes that maximize an expected information gain. In such a case the design variables would be the positions of the probes. Or one might want to optimize the physical scenario itself, e.g. find the best shape of an object that is part of the validation experiment. In such a case the design variables would be the parameters that govern the shape of the object. Or yet, all that can be done is to select the best shape among a given finite amount of available shapes. In such a case the design variable would be just an index assuming a value for each possible shape.

6. Virtual Experiments for Multiscale Models

In the case of multiscale models such as that discussed in Section 2, we may have access to additional prior information (often overlooked) available on the fabrication process (or, more generally, on microscale physical events) leading to the primal model being subjected to validation processes. The model of the forward problem in nonlinear elastostatics embodied

in (8), for example, is known to be the result of averaging events at a molecular scale that described polymerization and densification steps in fabrication.

Returning to the example of the parametric class of multiscale models (8), we now consider "virtual" calibration and validation scenarios based on the molecular substructure produced in the polymerization step in fabrication.

6.1. Virtual Validation Processes

The full process for the sample class of problems in Section 2 is as follows:

1. *Virtual Calibration Priors.* In the absence of information on the macroscale parameters $\boldsymbol{\theta} = (\alpha, \beta, \gamma)$, virtual priors can be calculated using (as "truth") the fine-scale molecular model. For example, a cubic representative volume element (RVE) can be generated for the molecular model, such as is shown in Figure 4, with prescribed axial, biaxial, and deformations mimicking laboratory experiments, for each of a set of $N_{\rm c} \geqslant 1$ realizations of the polymerization. The RVE volume is increased incrementally until results stabilize near a statically meaningful state. For each prescribed deformation pattern, the total energy, $V(\boldsymbol{\mu}; \mathbf{u}^n)$, of the molecular model is computed, so that $E/$volume is an approximation of the stored energy function W of (5), yielding for three independent virtual tests, and three equations for the three unknown parameters (α, β, γ). Repeating this process for the $N_{\rm c}$ realizations produces histograms for the variation of α, β, and γ, which approximate the joint pdf
$$\pi^{\rm virtual}(\boldsymbol{\theta}) \approx \pi(\boldsymbol{\theta}), \tag{45}$$
where $\pi(\boldsymbol{\theta})$ denotes the theoretical prior pdf defined via all possible realizations of the polymerization. Calculations described in [8] show that reasonable representations of parameter uncertainties can be obtained through this process.

2. *Macroscale Calibration.* Actual laboratory experiments are next run on physical material specimens of the elastomer involving calibration data $\mathbf{d}_{\rm c}$. The calibrated posterior is then
$$\pi(\boldsymbol{\theta}|\mathbf{d}_{\rm c}) = \frac{\pi(\mathbf{d}_{\rm c}|\boldsymbol{\theta})\ \pi^{\rm virtual}(\boldsymbol{\theta})}{\pi(\mathbf{d}_{\rm c})}. \tag{46}$$

3. *Virtual Validation.* Consider a possible validation scenario $\hat{s}_{\rm v}$ such as the prismatic molecular domain shown in Figure 5a. The top planar boundary, denoted Γ, is formed by the points $(x, y, z) \in \mathbb{R}^3$ that have $z = 0$ and are also located inside the perimeter defined by the the corner points A, B, C, and D. The bottom planar boundary has points with $z = c$. On Γ we prescribe a linear varying normal pressure field
$$\begin{aligned} \mathbf{p}(x,y) \quad = \quad & \Big[p(x_A, y_A, 0) + \frac{x}{a} p(x_B, y_B, 0) + \\ & + \frac{y}{b} p(x_D, y_D, 0) + \frac{xy}{ab} p(x_C, y_C, 0)\Big]\, \mathbf{n}_0, \end{aligned} \tag{47}$$
where $\mathbf{n}_0$ is a unit vector normal to Γ, and $p(x_\xi, y_\xi, 0)$, $\xi = A, B, C, D$, are the normal pressure magnitudes prescribed at points A, B, C, and D. We then set the normal force at any particle site $(x_i, y_i, 0) \in \Gamma$ to be given by
$$\mathbf{f}_i = \mathbf{p}(x_i, y_i) \frac{A_\Gamma}{N_\Gamma}, \tag{48}$$

where A_Γ is the area of Γ and N_Γ is the number of particles on Γ. In agreement with (2), these particle forces are understood to be given by

$$\mathbf{f}_i = \frac{\partial V(\boldsymbol{\mu}; \mathbf{u}^n)}{\partial \mathbf{u}_i}, \quad i \in I_\Gamma \tag{49}$$

where $\mathbf{U} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N\}$ is the vector of particle site displacements for the molecular statics problem (2), and I_Γ is the index set of indices of particles on Γ.

We next take the observables to be the corner displacements Δ_A, Δ_B, Δ_C, and Δ_D, normal to Γ and computed with the molecular model (see Subsection 6.2). The virtual validation data $\hat{\mathbf{d}}_\mathrm{v}$ is then defined as

$$\hat{\mathbf{d}}_\mathrm{v} = (\Delta_A, \Delta_B, \Delta_C, \Delta_D) \tag{50}$$

for each polymerization realization, from a total of $N_\mathrm{v} \geqslant 1$ realizations, as explained in Subsection 6.2. An example of such a virtual scenario domain $\Omega(\hat{s}_\mathrm{v})$ is shown in Figure 6. This virtual scenario should also provide an opportunity to check the validity of basic hypotheses used to define the molecular model, by allowing one to explore the extent of homogeneity and isotropy in results over several realizations.

Turning now to the continuum model being validated, the equilibrium states of the body are governed by minimizing the functional

$$\mathcal{F}(\mathbf{v}) = \int_{\Omega_0(\hat{s}_\mathrm{v})} W[(\alpha, \beta, \gamma); \mathbf{v}] \; d\Omega - \int_\Gamma \; \mathbf{t}_\Gamma \cdot \mathbf{n}_0 v_\Gamma \; d\Gamma, \tag{51}$$

where $\mathbf{v}$ is a trial function in an appropriate space $\mathcal{V}(\Omega_0(\hat{s}_\mathrm{v}))$ of trial functions, $W = W[(\alpha, \beta, \gamma); \mathbf{v}]$ is the stored energy function (5), $\mathbf{t}_\Gamma$ is the applied traction field normal to the surface Γ, and v_Γ is the normal trace of $\mathbf{v}$ on Γ. The minimizers $\mathbf{u} \in \mathcal{V}(\Omega_0(\hat{s}_\mathrm{v}))$ of (51) satisfy (8) and the traction boundary condition $\mathbf{P} \cdot \mathbf{n}_0 = \mathbf{t}_\Gamma$ on Γ:

$$\left.\begin{array}{rcll} \boldsymbol{\nabla} \cdot \dfrac{\partial W[(\alpha, \beta, \gamma); \mathbf{u}]}{\partial \mathbf{F}} & = & \mathbf{0} & \text{on } \Omega_0(\hat{s}_\mathrm{v})), \\ \mathbf{P} \cdot \mathbf{n}_0 & = & \mathbf{t}_\Gamma & \text{on } \Gamma, \\ \mathbf{u}(x, y, c) & = & 0, & \end{array}\right\} \tag{52}$$

where $\mathbf{P} = \partial W / \partial \mathbf{F}$ is the first Piola-Kirchnoff stress tensor, and $\mathbf{F}$ is the deformation tensor (3).

In general, only numerical approximations of solutions of (52) are obtainable; so, we cover $\Omega_0(\hat{s}_\mathrm{v})$ with a finite element mesh, as suggested by Figure 5b, and compute finite element approximations of the solution of (52) for fixed $\boldsymbol{\theta} = (\alpha, \beta, \gamma)$ following long-established methods [42]. We connect the virtual (molecular) model with the discrete approximation of the continuum model by setting

$$\mathbf{t}_\Gamma = \mathbf{p}(x, y). \tag{53}$$

With this boundary data now prescribed, we can solve the discrete model of (52) and compute the model outputs $\hat{\mathbf{y}}_\mathrm{v} = \hat{\mathbf{y}}_\mathrm{v}(\mathbf{u}_h(\boldsymbol{\theta}, \hat{s}_\mathrm{v}))$ for $\hat{\mathbf{d}}_\mathrm{v}$,

$$\hat{\mathbf{y}}_\mathrm{v} = \left[\begin{array}{l} \mathbf{n}_0 \cdot \mathbf{u}_h(\boldsymbol{\theta}, \hat{s}_\mathrm{v}; A) \\ \mathbf{n}_0 \cdot \mathbf{u}_h(\boldsymbol{\theta}, \hat{s}_\mathrm{v}; B) \\ \mathbf{n}_0 \cdot \mathbf{u}_h(\boldsymbol{\theta}, \hat{s}_\mathrm{v}; C) \\ \mathbf{n}_0 \cdot \mathbf{u}_h(\boldsymbol{\theta}, \hat{s}_\mathrm{v}; D) \end{array}\right], \tag{54}$$

for any $\boldsymbol{\theta}$ in the support of the posterior pdf (46). In (54), $\mathbf{u}_h(\boldsymbol{\theta}, \hat{s}_{\rm v}; \xi)$, $\xi = A, B, C, D$, are obviously the corner values of the finite element approximation $\mathbf{u}_h$ on Γ.

Comparing the discrete model outputs (54) with the virtual data (50), and assuming additive data noise and modeling error, we have

$$\hat{\mathbf{y}}_{\rm v} - \hat{\mathbf{d}}_{\rm v} = \boldsymbol{\varepsilon}. \tag{55}$$

The uncertainty embodied in $\boldsymbol{\varepsilon}$ reflects assumptions about the discrepancy between observations and model outputs, and it is not to be confused with ε_s of (21): here $\varepsilon = -(\mathbf{b}_s + \varepsilon_s)$. The discrepancy model can itself be subjected to calibration, in which case the random vector $\boldsymbol{\theta}$ is augmented with extra parameters. Customarily, in statistical inverse problems (see e.g. [61]), one assumes a Gaussian $\boldsymbol{\varepsilon}$ with zero mean and covariance matrix $\mathbf{C}_\varepsilon$, causing the likelihood of $\hat{\mathbf{d}}_{\rm v}$ to have the form

$$\pi(\hat{\mathbf{d}}_{\rm v}|\boldsymbol{\theta}, \mathbf{d}_{\rm c}) \propto \exp\left\{-\frac{1}{2}[\hat{\mathbf{y}}_{\rm v} - \hat{\mathbf{d}}_{\rm v}]^T \cdot \mathbf{C}_\varepsilon^{-1} \cdot [\hat{\mathbf{y}}_{\rm v} - \hat{\mathbf{d}}_{\rm v}]\right\}. \tag{56}$$

Thus, the virtual validation experiment leads to the posterior pdf,

$$\pi(\boldsymbol{\theta}|\hat{\mathbf{d}}_{\rm v}, \mathbf{d}_{\rm c}) = \frac{\pi(\hat{\mathbf{d}}_{\rm v}|\boldsymbol{\theta}, \mathbf{d}_{\rm c})\pi(\boldsymbol{\theta}|\mathbf{d}_{\rm c})}{\pi(\hat{\mathbf{d}}_{\rm v}|\mathbf{d}_{\rm c})}. \tag{57}$$

This process can, of course, be repeated for other virtual validation scenarios, such as those pictured symbolically in Figure 7. A more general approach is outline in Section 7 that targets estimating model bias.

4. *Information Gain.* It remains to choose the validation scenario that produces the largest expected information gain over the calibration posterior pdf (46), among all scenarios considered. The virtual validation scenarios leading to the largest information gain are judged to be candidates of the construction and implementation of actual, physical validation experiments.

5. *Validation.* At this point, one may proceed to construct actual validation experiments approximated by the virtual scenarios and validation calculations. The remainder of the process leading to the validation test (36) is then completed.

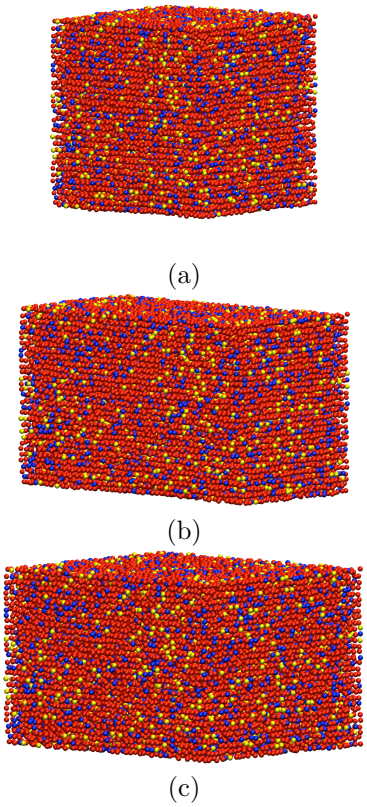
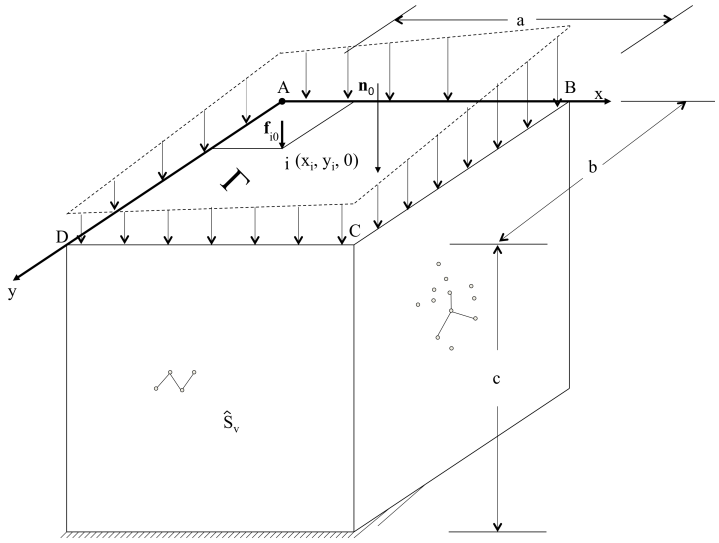


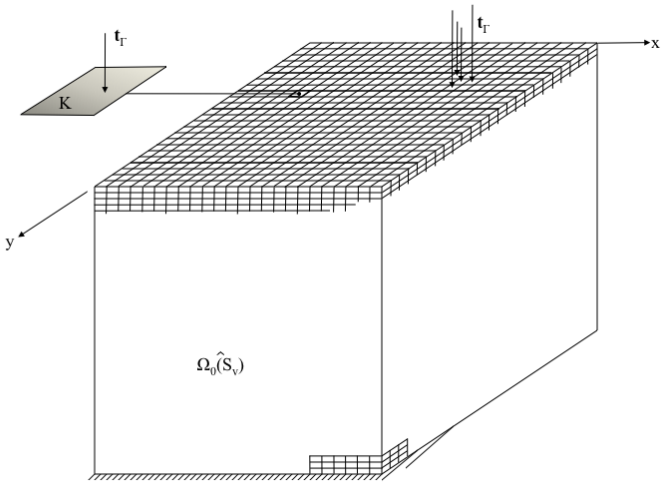
Figure 4: Molecular models of RVEs used in the calibration phase to produce priors for macroscale model parameters (taken from [9]): a) relaxation, b) uniaxial stretch, and c) biaxial stretch.

6.2. Solution of the Molecular Model

To obtain numerical solutions to the molecular model of the validation experiments, we employ the methods of adaptive multiscale modeling developed in [8, 9, 51]. The idea is to replace the molecular model with a hybrid multiscale model in which a continuum-hyperelastic model based on (8) is used in the part of the domain $\Omega_0(\hat{s}_{\rm v})$, and a molecular model is used in the neighborhood of data of interest (in the earlier example, those would be the neighborhood of corners of the plane Γ), and an interface between the molecular model and a finite-element approximation of the continuum model is implemented using the Arlequin method. By computing a posteriori estimates of errors in the data of interest, relative to the fine-scale molecular model, an adaptive modeling scheme can be used to furnish fine-scale information sufficient to control the error in the data of interest. Interestingly, it is sufficient to use calibrated pdfs for the continuum model parameters $\boldsymbol{\theta}$ in these algorithms.



(a)



(b)

Figure 5: (a) Molecular model of a validation scenario subject to prescribed discrete force field on Γ and probed corner displacements Δ_A, Δ_B, Δ_C, and Δ_D, and (b) a finite element discretization of the continuum model of the scenario subjected to surface tractions $\mathbf{t}_\Gamma$.

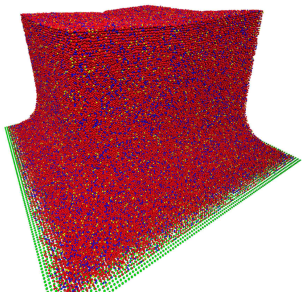


Figure 6: Three-million degree of freedom molecular model of one realization of the problem domain $\hat{s}_{\rm v}$ (after [9]).

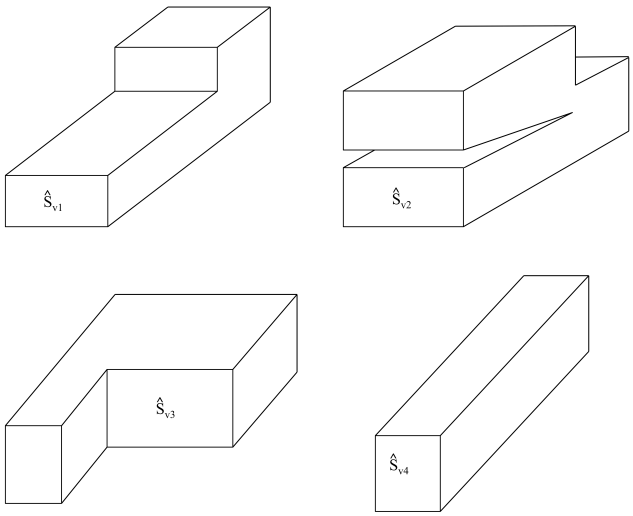


Figure 7: Symbolic view of sequence of possible validation scenarios.

Let the vector $\hat{\mathbf{d}}_{\rm v} \in \mathbb{R}^{n_d}$ designate the data to be generated by the molecular model, and let $\hat{d}_\ell$ designate the ℓ-th component of such vector, $\ell = 1, \ldots, n_d$. The algorithmic steps in this process are summarized as follows:

1. Given an integer $N_{\rm v} \geqslant 1$, perform steps 2-8 for $r = 1, 2, \ldots, N_{\rm v}$.

2. A polymeric densification model of the domain $\hat{s}_{\rm v}$ (plus boundary conditions) is generated following the process described in Subsection 6.1, with lattice-site potentials E_{ik} of (2) known for the r-th realization of the molecular structure. The fine scale forward problem is then defined by the system (2). A weak form of the forward problem is:

   Find $\mathbf{U} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N\} \in \mathcal{U}$ such that

$$R(\boldsymbol{\mu}; \mathbf{U}; \mathbf{V}) = 0 \qquad \forall\ \mathbf{V} \in \mathcal{V}, \tag{58}$$

where $R(\cdot;\cdot)$ is the residual functional,

$$R(\boldsymbol{\mu};\mathbf{U};\mathbf{V}) = \sum_{i=1}^{N} \left\{ \frac{\partial V(\boldsymbol{\mu};\mathbf{u}^n)}{\partial \mathbf{u}_i} - \mathbf{f}_i \right\} \cdot \mathbf{v}_i, \tag{59}$$

and $\mathcal{U}$ and $\mathcal{V}$ are appropriate trial and test spaces of N vectors in $\mathbb{R}^3$, $\mathcal{V}$ being the space of test functions satisfying appropriate boundary conditions. Throughout, we assume the fine-scale model of the polymer densification is (virtually) a representation of physical reality; that is, the parameters in the definitions of molecular potentials are known exactly, even though each realization of the polymerization leads to different conformations of the molecular constituents.

3. Construct the adjoint problem to (58),

$$\frac{\partial R(\boldsymbol{\mu};\mathbf{U};\mathbf{V})}{\partial \mathbf{U}} : \mathbf{P} = \hat{d}_\ell(\mathbf{V}) \qquad \forall\ \mathbf{V} \in \mathcal{V} \tag{60}$$

or, equivalently,

$$\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial^2 V(\boldsymbol{\mu};\mathbf{u}^n)}{\partial \mathbf{u}_i \partial \mathbf{u}_j} \mathbf{p}_j^{(\ell)} \cdot \mathbf{v}_i = \hat{d}_\ell(\mathbf{V}) \qquad \forall\ \mathbf{V} \in \mathcal{V}, \tag{61}$$

where $\mathbf{p}^{(\ell)} = \{\mathbf{p}_1^{(\ell)}, \mathbf{p}_2^{(\ell)}, \ldots, \mathbf{p}_N^{(\ell)}\}$ is the solution to the adjoint problem, given the solution $\mathbf{U}$ to (58), and $\hat{d}_\ell$ is the data to be generated by the molecular model, as discussed in Subsection 6.1. Here we assume $\hat{d}_\ell$ is characterized by a linear functional on $\mathcal{V}$, but nonlinear generalizations are straightforward [51].

4. Construct a hybrid-continuum-molecular model, with molecular structure near the subdomain containing the data to be measured, continuum (in this case, an elastomer modeled using the system (8)) with an appropriate molecular-continuum interface Γ implemented using the Arlequin method (see [8, 9]).

5. Solve the hybrid model for the hybrid solution $\widehat{\mathbf{U}}^h$ and the corresponding hybrid model approximation $\widehat{\mathbf{d}}_{\mathrm{v}}^h = (\hat{d}_1^h, \ldots, \hat{d}_{n_d}^h)$ to the data $\widehat{\mathbf{d}}_{\mathrm{v}}^{(r)} = (\hat{d}_1, \ldots, \hat{d}_{n_d})$.

6. To first order [9, 15, 49, 51], the error in the hybrid model approximation of $\widehat{\mathbf{d}}_{\mathrm{v}}^{(r)}$ is

$$\varepsilon_\ell \equiv |\hat{d}_\ell - \hat{d}_\ell^h| \approx |R(\widehat{\mathbf{U}}^h, \mathbf{p}^{(\ell)})|, \tag{62}$$

where $\mathbf{p}^{(\ell)}$ is the solution to the adjoint problem (60). Let us define

$$\epsilon = \max\{\varepsilon_1, \ldots, \varepsilon_{n_d}\}. \tag{63}$$

7. Introduce an error tolerance ϵ_{tol} for the modeling error ϵ of (63) and implement the goal-oriented adaptivity algorithm of [9, 15]. If $\hat{\epsilon} > \epsilon_{\mathrm{tol}}$, more fine-scale information is added to the hybrid model. An example of steps in the adaptive modeling algorithm is illustrated in Figure 8.

8. Continue the adaptive modeling process until $\epsilon \leq \epsilon_{\mathrm{tol}}$.

Steps 2-8 are performed for N_{v} realizations of the polymerization molecular structure, allowing the generation of N_{v} samples of the random variable $\widehat{\mathbf{D}}_{\mathrm{v}}$.

This process leads to an approximation of the random validation data $\widehat{\mathbf{D}}_{\mathrm{v}}$ and its uncertainty due to multiple realizations of the polymerization process.

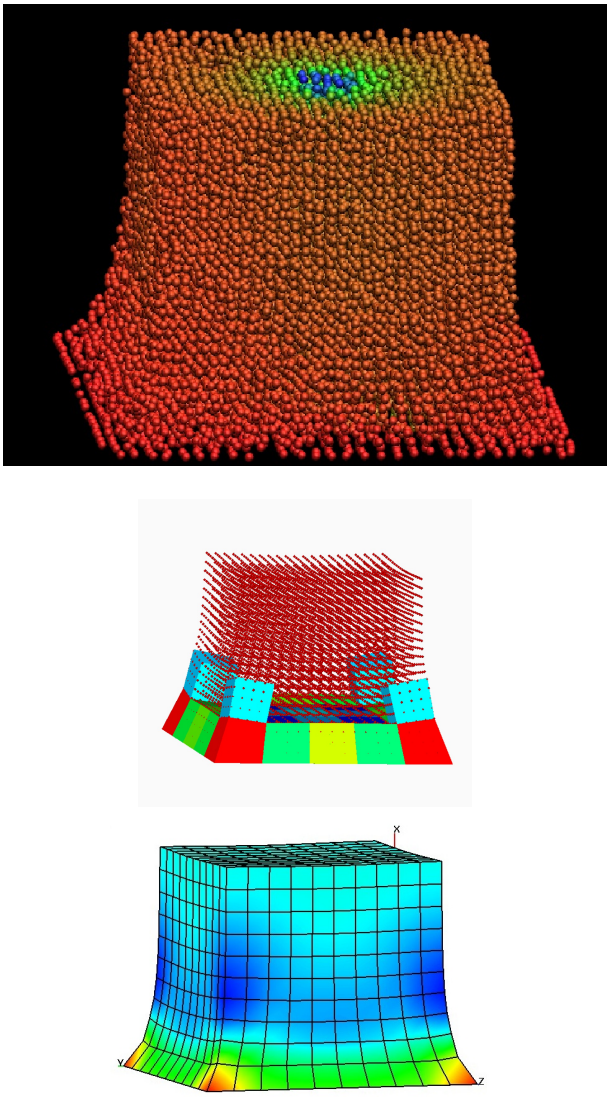


Figure 8: A sequence of adaptive models of a possible validation scenario, beginning with the full molecular model (top subfigure) and adaptively adding data (the other two subfigures) to a continuous hybrid model to control the error in quantities of interest, such as validation data (adapted from [9]).

7. Approximation of Model Bias and Hyper-Model Plausibilities

7.1. Virtual Calibration

Let us return to the example described in Section 6 and consider the problem of calibration and of determining the model bias in models such as that described by (22) and (23). We consider a calibration in which a coupon of polymer material is identified which forms a prismatic

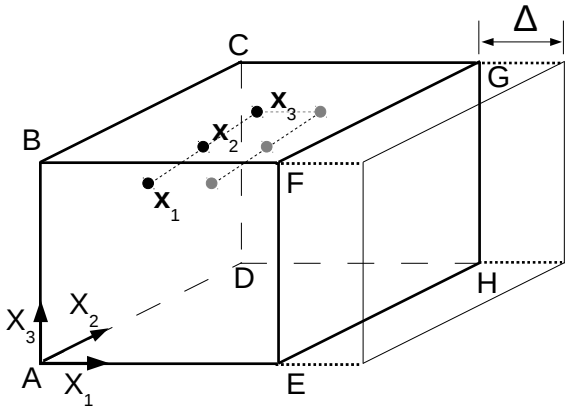


Figure 9: A realization of a molecular model of a calibration domain Ω subjected to controllable kinematic constraints, and probe points $\mathbf{x}_i \in \Omega,\ i = 1, 2, \ldots, N$.

RVE occupying a region Ω in $\mathbb{R}^3$, as illustrated in Fig. 9. This is one realization of a molecular-scale calibration scenario. As one generates polymer conformations through polymerization algorithms, generally driven by kinetic Monte Carlo strategies, many consistent realizations of the polymer molecular structure are generated. Experiments (virtually) on RVEs for each realization yield histograms which define the uncertainty in key parameters once the molecular model parameters are known. In each such experiment, we prescribe a controlled kinematic constraint defined by the extension of plane EFGH an amount Δ in the x_1-direction, while keeping plane $ABCD$ planar and fixed. Let $\mathbf{q} = \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ denote a $3n$-vector defining a QoI:

$$q(\mathbf{v}) = \sum_{i=1}^{n} \mathbf{q}_i \cdot \mathbf{v}_i, \quad \mathbf{v} \in \mathcal{U}. \tag{64}$$

In this case, the QoI is an output outside the set of observational data vector $\mathbf{d}_{\mathrm{c}}$.

Let $\mathbf{x}_I^*$ denote a particular molecule site in domain Ω_{c} of s_{c} at which the solution $\mathbf{u}_*^n$ will be sampled; *i.e.* $\mathbf{x}_I^*$ is an input point, and $\mathbf{u}_I^*$ the displacement of molecule I. Set $\mathbf{q} = \{\mathbf{0}, \mathbf{0}, \ldots, \mathbf{1}, \ldots, \mathbf{0}\}$, where $\mathbf{1}$ is a unit vector at site $\mathbf{x}_I^*$ in the direction of the observational probe. Then

$$q(\mathbf{v}) = \mathbf{1} \cdot \mathbf{v}_I. \tag{65}$$

Let $\mathbf{p}_*$ denote the solution of the adjoint problem (61) for this choice of a QoI. Then the "experimental" observation of the displacement component u_I^* is $q(\mathbf{u}_*^n) = u_I^*$.

Next, let $\mathbf{s}^n$ be an arbitrary $3n$-vector in U. It is proved in [49] (see also [51]) that

$$q(\mathbf{u}_*^n) - q(\mathbf{s}^n) = R(\boldsymbol{\mu}; \mathbf{s}^n, \mathbf{p}_*) + \Delta, \tag{66}$$

where $R(\cdot; \cdot, \cdot)$ is the residual (59) and Δ is a remainder of order $\|\mathbf{u}_*^n - \mathbf{s}^n\|^\sigma$, $\sigma \geq 2$, $\|\cdot\|$ being the Euclidean norm on $\mathcal{U}$. In general, Δ is neglected if $\mathbf{s}^n$ is "close" to $\mathbf{u}_*^n$ (recall (62)). Then $R(\boldsymbol{\mu}; \mathbf{u}_*^n, \mathbf{p}_*)$ provides an *a posteriori* estimate of the error produced by approximating $q(\mathbf{u}_*^n)$ by $q(\mathbf{s}^n) = \mathbf{1} \cdot \mathbf{s}_I^*$.

Let us now suppose that we cover the domain Ω_c with a finite element mesh Ω_c^h with N^h nodes, each coinciding with a molecule site (generally, $N^h \ll n$) and let $\mathbf{w}^h$ denote a finite element approximation of the solution of the continuum-macromodel $\mathbb{M}$ for the s_c geometry and boundary conditions. Let $\Pi\mathbf{w}^h$ denote a projection of $\mathbf{w}^h = \mathbf{w}^h(\boldsymbol{\theta})$ onto the molecule sites, so $\Pi\mathbf{w}^h \in \mathcal{U}$. Here $\boldsymbol{\theta}$ designates the parameters of the macromodel $\mathbb{M}$. Denote $q(\Pi\mathbf{w}^h) = w_I^{h*}$. Then (66) reduces to

$$\mathbf{1} \cdot \mathbf{u}_I^* - w_I^{h*} = R(\boldsymbol{\mu}; \Pi\mathbf{w}^h, \mathbf{p}_*). \tag{67}$$

The interpretation of (67) is as follows. The macroscale model of equilibrium deformation of an elastomeric body, our model problem, is replaced by a computational model defined by a finite element approximation, yielding a solution $\mathbf{w}^h(\boldsymbol{\theta}; \mathbf{x})$, $\mathbf{x} \in \Omega_c$. For mesh size h sufficiently small, $\mathbf{w}^h$ is assumed to be a good approximation of $\mathbf{w}$. We probe model outputs at controllable input point $\mathbf{x}_I^*$ which also are molecule sites in the fine-scale molecular model of the system: $\Pi\mathbf{w}^h(\boldsymbol{\theta}, \mathbf{x}_I^*) = \mathbf{w}_I^{h*}$. The solution of the molecular model is $\mathbf{u}_*^n$ and its depiction of the coordinates of the displaced input point $\mathbf{x}_I^*$ is $\mathbf{r}_I^*$. Thus (67) represents a fine-scale approximation of the modeling error, the model bias $\mathbf{b}(\mathbf{x}_I^*) = \mathbf{b}_I^*$ at input point $\mathbf{x}_I^*$: $b_I^* = \mathbf{1} \cdot \mathbf{b}_I^*$,

$$b_I^* = R(\boldsymbol{\mu}; \Pi\mathbf{w}^h, \mathbf{p}_*). \tag{68}$$

We repeat this calculation at m points $\mathbf{x}_k^* \in \Omega_c$, $k = 1, 2, \ldots, m$ to obtain a discrete characterization of model bias $\mathbf{b}_I$ over Ω_c for the calibration scenario.

We now repeat this process for many realizations of the polymer molecular structure, producing a histogram of each bias component (68) which approximates a pdf $\pi_{\mathbf{b}}(b_I^*)$. The collection of such pdfs $\pi_{\mathbf{b}}(b_I^*)$, for various molecular sites I, allows us to produce a fine-scale approximation of the likelihood probability $\pi(\mathbf{d}|\boldsymbol{\theta}, \mathbb{M})$ of a data set $\mathbf{d}$ generated by the molecular model, conditioned on the macromodel $\mathbb{M}$ and its parameters $\boldsymbol{\theta}$.

7.2. Other Bias Models

Procedures that attempt to account for model bias generally involve introducing a new uncertainty model for model discrepancy, ranging from the common strategy of assuming simply at Gaussian likelihood with zero mean as done in (56) and often in the literature (*e.g.* [28, 29, 45, 61]), to the use of Gaussian processes to depict differences between observational data (corrected for experimental noise) and model predictions [33, 10], to the use of embedded models for model error [38], to the use of fine-scale data as described in the development in Subsection 7.1. In all cases, the introduction of such models of bias involves the addition of new parameters into the process; *e.g.* the hyperparameters of Gaussian processes or the fine-scale model parameters $\boldsymbol{\mu}$. These additional parameters may be determined from prior information available on the underlying fine-scale model (such as the case when the interaction potential parameters $\boldsymbol{\mu}$ are known) or they may be added to the model parameters $\boldsymbol{\theta}$ to be updated in the Bayesian calibration and validation processes, or, in some cases, as in [10, 25], they are determined approximately using MLE (Maximum Likelihood) methods.

It would seem that for models of physical phenomena such as those considered in this work, the most satisfactory approach would be to generate models of bias based on what is known about the purposes of the problem at hand. Beginning with the elastostatics model characterized by (5) and (7) (or by (8)), one could consider more general hierarchical families of models such as those for which

$$\mathbf{P} = \frac{\partial}{\partial \mathbf{F}} \left\{ \sum_{r=1}^{m_1} \sum_{s=1}^{m_2} \sum_{t=1}^{m_3} C_{rst} (I_1(\mathbf{C}) - 1)^r (I_2(\mathbf{C}) - 1)^s (J(\mathbf{C}) - 1)^t - \kappa \ln J(\mathbf{C}) \right\} + \zeta \mathbf{G}(\dot{\mathbf{C}}), \tag{69}$$

where $\mathbf{P}$ is the Piola-Kirchhoff stress (recall (52)) and now the model parameters are the coefficients $\{C_{rst}\}$ and ζ and $\mathbf{G}$ is a factor defining rate effects. Various possible improvements in the original model are obtained by truncating the series in (69). Thus, we can generate a family $\mathcal{M}$ of models representing possible improvements of the original model, each with the original parameters $\boldsymbol{\theta} = (\alpha, \beta, \gamma)$ plus additional paramters as in (69).

Let $\mathbb{M}_k \in \mathcal{M}$ be a model in this family, and let $\boldsymbol{\theta}_k$ be its parameters. The posterior plausibility of model $\mathbb{M}_k \in \mathcal{M}$ is defined by

$$\rho(\mathbb{M}_k|\mathbf{d}, \mathcal{M}) = \frac{\pi(\mathbf{d}|\mathbb{M}_k)\rho(\mathbb{M}_k|\mathcal{M})}{\pi(\mathbf{d}|\mathcal{M})}, \tag{70}$$

where $\mathbf{d}$ is the data produced by the molecular model, $\pi(\mathbf{d}|\mathbb{M}_k)$ is the evidence of the macro-model $\mathbb{M}_k$,

$$\pi(\mathbf{d}|\mathbb{M}_k) = \int \pi(\mathbf{d}|\boldsymbol{\theta}_k, \mathbb{M}_k)\pi(\boldsymbol{\theta}_k|\mathbb{M}_k)\, d\boldsymbol{\theta}_k, \tag{71}$$

and $\rho(\mathbb{M}_k|\mathcal{M})$ is a prescribed prior plausibility of $\mathbb{M}_k$ in $\mathcal{M}$. We note that

$$\sum_k \rho(\mathbb{M}_k|\mathbf{d}, \mathcal{M}) = 1. \tag{72}$$

Thus, the plausibilities provide a powerful tool to judge which models among the set $\mathcal{M}$ are the most plausible given the data, the models with plausibilities closest to unity being the most plausible and the best to reduce the bias in the original model.

The plausibility ρ computed as in (70) assumes that the parameters $\boldsymbol{\mu}$ of the molecular model, used to produce the data $\mathbf{d}$, are known exactly, which is not generally the case. More generically, one has

$$\pi(\mathbf{d}|\mathbb{M}_k) = \int_{\boldsymbol{\mu}\ \text{space}} \pi(\mathbf{d}|\boldsymbol{\mu}, \mathbb{M}_k)\pi(\boldsymbol{\mu}|\mathbb{M}_k)\, d\boldsymbol{\mu}. \tag{73}$$

The integral (73) can be approximated by a Monte Carlo procedure,

$$\pi(\mathbf{d}|\mathbb{M}_k) \approx \frac{1}{J}\sum_{j=1}^{J} \pi(\mathbf{d}|\boldsymbol{\mu}^{(j)}, \mathbb{M}_k), \tag{74}$$

where the J samples $\boldsymbol{\mu}^{(j)},\ j = 1, 2, \ldots, J$, are generated according to the given distribution $\pi(\boldsymbol{\mu}|\mathbb{M}_k)$. Usually, the distribution of $\boldsymbol{\mu}$ can be assumed to be independent of $\mathbb{M}_k$.

8. Other Remarks

An issue that arises is what provisions need to be made for cases in which a vector $\mathbf{Q} = \{Q_1, Q_2, \ldots, Q_{n_q}\}$ of multiple QoIs is designated in the prediction scenario. The answer is that the processes developed in this study carry over without modification to problems with multiple QoIs. The validation criteria may be applied to each Q_i using the same process, in some cases possibly without additional validation experiments. In such cases, one may find that the validation criterion is satisfied for some QoIs but not all, in which case the model is not invalid for those passing the test (36) but invalid for the remaining QoIs. It is even possible to apply the same processes to the joint pdf of all components in $\mathbf{Q}$, or to joint pdfs of subgroups of $\mathbf{Q}$ components.

Finally, we remark that the idea of using microscale models as a basis for the design of validation experiments need not be restricted to the very complex molecular models of the type described here. Any reasonable model that results from homogenization processes has as its base a fine-scale structure that could, in theory, be used to supply additional information for the design of experiments.

Acknowledgements

This work was partially supported by the Mathematical Multifaceted Integrated Capability Centers (MMICCs) effort within the Applied Mathematics activity of the U.S. Department of Energy's Advanced Scientific Computing Research program under Award Number DE-5C0009286. We also gratefully acknowledge the support of this work under the DOE contract DE-FG02-05ER25701 in connection with the DOE Multiscale Mathematics Program and the DOE contract DE-FC52-08NA28615 in connection with the Predictive Science Academic Alliance Program. We thank Serge Prudhomme for comments on early drafts of this work and Ivo Babuska for critiquing and encouraging comments on this work as well. Kathryn Farrell kindly assisted typing various drafts of this work.

References

[1] AIAA, Guide for the verification and validation of computational fluid dynamics simulations (1998).

[2] S. Arlot, A. Celisse, A survey of cross-validation procedures for model selection, Statistical Surveys 4 (2010) 40–79.

[3] ASME Committee PTC-60 V & V 10, Guide for Verification and Validation in Computational Solid Mechanics, ASME, 2006.

[4] I. Babuška, J. T. Oden, Verification and validation in computational engineering and science: Part 1, basic concepts, Computer Methods in Applied Mechanics and Engineering 193 (1) (2004) 4047–4068.

[5] I. Babuška, J. T. Oden, The reliability of computer predictions: Can they be trusted?", International Journal of Numerical Analysis and Modeling 1 (1) (2005) 1–18.

[6] I. Babuška, R. Tempone, F. Nobile, A systematic approach to model validation based on Bayesian updates and prediction-related rejection criteria, Computer Methods in Applied Mechanics and Engineering 197 (2008) 2517–2539.

[7] R. Batra, Elements of Continuum Mechanics, AIAA Education Series, Reston, VA, 2006.

[8] P. T. Bauman, Adaptive multiscale modeling of polymeric materials using goal-oriented error estimation, arlequin coupling, and goals algorithms, Ph.D. thesis, The University of Texas at Austin (2008).

[9] P. T. Bauman, J. T. Oden, S. Prudhomme, Adaptive multiscale modeling of polymeric materials with arlequin coupling and goals algorithms, Computer Methods in Applied Mechanics and Engineering 193 (2009) 799–838.

[10] M. J. Bayarri, J. O. Berger, R. Paulo, J. Sacks, J. M. Cafeo, C. Cavendish, L. C.-H., J. Tu, A framework for validation of computer models, Technometrics 49 (2) (2007) 138–154.

[11] H. J. C. Berendsen, D. van der Spoel, R. van Drunen, Gromacs: A message-passing parallel molecular dynamics implementation, Computer Physics Communications 91 (13) (1995) 43 – 56.
URL http://www.sciencedirect.com/science/article/pii/001046559500042E

[12] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, CHARMM: A program for macromolecular energy, minimization and dynamics calculations, Journal of Computational Chemistry 4 (2) (1983) 187–217.

[13] R. L. Burns, S. C. Johnson, G. H. Schmid, E. K. Kim, M. D. Dickey, J. M. Meiring, S. D. Burns, N. A. Starey, G. C. Willson, Mesoscale modeling for sfil simulating polymerization kinetics and densification, Proceedings, SPIE 5374 (2004) 349–360.

[14] S. H. Cheung, T. A. Oliver, E. E. Prudencio, S. Prudhomme, R. D. Moser, Bayesian uncertainty analysis with applications to turbulence modeling, Reliability Engineering and System Safety 96 (2011) 1137–1149.

[15] P. G. Ciarlet, Mathematical Elasticity, Volume 1: Three-Dimensional Elasticity, No. v. 1 in Studies in Mathematics and Its Applications, North-Holland, 1998.

[16] M. A. Clyde, Experimental design: A bayesian perspective, International Encyclopedia of Social and Behavorial Sciences April (2001) 1–21.

[17] Committee on Mathematical Foundations of Verification, Validation, and Uncertainty Quantification; Board on Mathematical Sciences and Their Applications, Division on Engineering and Physical Sciences, National Research Council, Assessing the Reliability of Complex Models: Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification, The National Academies Press, 2012, http://www.nap.edu/openbook.php?record_id=13395.

[18] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, Journal of the American Chemical Society 117 (19) (1995) 5179–5197.

[19] T. M. Cover, J. A. Thomas, Elements of Information Theory, 2nd ed., Wiley - Interscience, Hoboken, 2006.

[20] K. Farrell, J. T. Oden, Statistical calibration and validation methods of coarse-grained and macro models of atomic systems, ICES REPORT 12-45 (2012) .

[21] S. Ferson, W. L. Oberkampf, L. Ginzburg, Model validation and predictive capability for the thermal challenge problem, Computer Methods in Applied Mechanics and Engineering 197 (2008) 2408–2436.

[22] L. Foglia, S. W. Mehl, M. C. Hill, P. Perona, P. Burlando, Testing alternative ground water models using cross-validation and other methods, Ground Water 45 (5) (2007) 627–642.

[23] M. E. Gurtin, An Introduction to Continuum Mechanics, Academic Press, New York, 1981.

[24] A. Hadler, R. Bhattacharya, Model validation: A probabilistic formulation, in: 50th IEEE Conference on Decision and Control and European Control Conference, Orlando, FL, 2011.

[25] D. Higdon, J. Gattiker, B. Williams, M. Rightley, Computer model calibration using high-dimensional output, Journal of the American Statistical Association 103 (2008) 570–583.

[26] C. Howson, P. Urbach, Scientific Reasoning: The Bayesian Approach, 3rd ed., Open Court, Chicago and La Salle, 2006.

[27] E. T. Jaynes, Theory of Probability: The Logic of Science, Cambridge University Press, Cambridge, U.K., 2003.

[28] X. Jiang, S. Mahadevan, Bayesian cross-entropy methodology for optimal design of validation experiments, Measurement Science and Technology 17 (2006) 1895–1908.

[29] X. Jiang, S. Mahadevan, Bayesian validation assessment of multivariate computational models, Journal of Applied Statistics 15 (1) (2008) 49–65.

[30] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, Journal of the American Chemical Society 118 (45) (1996) 11225–11236.

[31] W. L. Jorgensen, J. Tirado-Rives, The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin, Journal of the American Chemical Society 110 (6) (1988) 1657–1666.

[32] M. C. Kennedy, A. O'Hagan, Predicting the output from a complex computer code when fast approximations are available, Biometrika 87 (2000) 1–13.

[33] M. C. Kennedy, A. O'Hagan, Bayesian calibration of computer models, Journal of the Royal Statistical Society. Series B (Statistical Methodology) 63 (3) (2001) pp. 425–464. URL http://www.jstor.org/stable/2680584

[34] O. P. Le Maitre, O. M. Knio, Spectral Methods for Uncertainty Quantification, With Applications to Computational Fluid Dynamics, Springer, 2010.

[35] A. R. Leach, Molecular Modeling: Principles and Applications, 2nd ed., Pearson Education Limited, Prentice Hall, Harlow, 2001.

[36] F. C. Mish, editor in chief, Merriam-Webster Collegiate Dictionary, 10th ed., Merriam-Webster, Inc., Springfield, 1993.

[37] M. J. Mooney, A theory of large elastic deformation, Journal of Applied Physics 11 (1940) 582.

[38] R. D. Moser, G. Terejanu, T. Oliver, C. Simmons, Validating and prediction of unobserved quantities, ICES REPORT 12-32 (2012) .

[39] J. I. Myung, M. A. Pitt, Optimal experimental design for model discriminization, Psychological Review 116 (2009) 499–518.

[40] W. L. Oberkampf, C. Roy, Verification and Validation in Scientific Computing, Cambridge University Press, Cambridge, UK, 2010.

[41] W. L. Oberkampf, T. G. Trucano, C. Hirsch, Verification, validation, and predictive capability in computational engineering and physics, Applied Mechanics Reviews 57 (2004) 345.

[42] J. T. Oden, Finite Elements of Nonlinear Continua, McGraw-Hill, N.Y, 1972. Dover Edition, 2006.

[43] J. T. Oden, An Introduction to Mathematical Modeling: A Course in Mechanics, Wiley Series in Computational Mechanics, John Wiley & Sons, 2011.

[44] J. T. Oden, A. Hawkins, S. Prudhomme, General diffuse-interface theories and an approach to predictive tumor growth modeling, Mathematical Models and Methods of Applied Science 20 (3) (2010) 1–41.

[45] J. T. Oden, R. Moser, O. Ghattas, Computer predictions with quantified uncertainty, "Part I", SIAM News 43 (10) (2010) .

[46] J. T. Oden, R. Moser, O. Ghattas, Computer predictions with quantified uncertainty, "Part II", SIAM News 43 (10) (2010) .

[47] J. T. Oden, E. E. Prudencio, Virtual model validation of complex multiscale systems: Applications to nonlinear elasticity, ICES REPORT 12-40.

[48] J. T. Oden, E. E. Prudencio, A. Hawkins-Daarud, Selection and assessment of phenomenological models of tumor growth, Mathematical Models and Methods in Applied Sciences 23 (2012) 1309–1338.

[49] J. T. Oden, S. Prudhomme, Estimates of modeling error in computational mechanics, Journal of Computational Physics 182 (2002) 496–515.

[50] J. T. Oden, S. Prudhomme, Control of modeling error in calibration and validation processes for predictive stochastic models, International Journal for Numerical Methods in Engineering (2010) 1–10.

[51] J. T. Oden, S. Prudhomme, A. Romkes, P. T. Bauman, Multiscale modeling of physical phenomena: adaptive control of models, SIAM Journal of Scientific Computing 28 (2008) 2359–2389.

[52] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, Journal of Computational Physics 117 (1995) 1–19.

[53] K. Popper, The Logic of Scientific Discovery, Routledge: Taylor & Francis, 1959.

[54] E. E. Prudencio, S. H. Cheung, Parallel adaptive multilevel sampling algorithms for the Bayesian analysis of mathematical models, International Journal for Uncertainty Quantification 2 (3) (2012) 215–237.

[55] E. E. Prudencio, K. W. Schulz, The Parallel C++ Statistical Library QUESO: Quantification of Uncertainty for Estimation, Simulation and Optimization, in: M. Alexander et al. (ed.), Euro-Par 2011 Workshops, Part I, vol. 7155 of Lecture Notes in Computer Science, Springer-Verlag, Berlin Heidelberg, 2012, pp. 398–407.

[56] R. S. Rivlin, Large elastic deformations of isotropic materials, Philosophical Transactions of the Royal Society 240 (822) (1948) 378.

[57] P. Roach, Verification and Validation in Computational Science and Engineering, Hermosa Press, Albuquerque, N.M., 1998.

[58] T. Santer, B. J. Williams, W. Notz, The Design and Analysis of Computer Experiments, Springer, 2003.

[59] C. E. Shannon, A mathematical theor of communication, Bell System Technical Journal 27 (1948) 379–423, 623–656.

[60] D. Sornette, A. B. Davis, K. Ide, K. R. Vixie, V. Pisarenko, J. R. Kamm, Algorithm for model validation: Theory and applications, PNAS 104 (16) (2007) 6562–6567.

[61] A. Tarantola, Inverse Problem Theory and Methods for Model Parameter Estimation, SIAM, Philadelphia, 2005.

[62] L. R. G. Trealor, The Physics of Rubber Elasticity, 2nd ed., Oxford University Press, 2006.

[63] A. F. Voter, Introduction to the Kinetic Monte Carlo Method, in: K. E. Sickafus, E. A. Kotomin (eds.), Radiation Effects in Solids, Lecture Notes in Computer Science, Springer NATO Publishing Unit, Dordrecht, The Netherlands, 2005.

[64] S. V. Weijs, R. van Nooijen, N. van de Geisen, Kullback-leibler divergence as a forecast skill score with classic reliability-resolution-uncertainty decomposition, in: Monthly Weather Review, American Meteorological Society, 2010, pp. 3387–3400.

[65] P. K. Weiner, P. A. Kollman, AMBER: Assisted Model Building with Energy Refinement. A general program for modeling molecules and their interactions, Journal of Computational Chemistry 2 (3) (1981) 287–303.

Appendix A. Molecular Models of Molecular Systems

Consider a molecular or atomic system consisting of n particles, in motion under the action of applied forces. A realization of the polymerization molecular structure and bonds forming polymer chains is illustrated in Fig. A.1. The state of the system is described by n coordinate-momentum pairs $\{(\mathbf{r}_i, \mathbf{p}_i)\}_{i=1}^n$, so that the instantaneous state of the system is defined by the set $\{\mathbf{r}_i, \mathbf{r}_2, \ldots, \mathbf{r}_n;\ \mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ written compactly as $(\mathbf{r}^n, \mathbf{p}^n)$. The Hamiltonian of such a system is of the form,

$$H(\mathbf{r}^n, \mathbf{p}^n) = \sum_{i=1}^{n} \frac{\mathbf{p}_i \cdot \mathbf{p}_i}{2m_i} + V(\mathbf{r}^n) - \sum_{i=1}^{n} \mathbf{f}_i \cdot \mathbf{u}_i, \tag{A.1}$$

where m_i is the mass of the particle (atom or moelcular) at site i, $V(\mathbf{r}^n)$ is the total potential energy, and $\mathbf{f}_i$ is the applied force at site i. The equations of motion are:

$$\dot{\mathbf{r}} = \frac{\partial H(\mathbf{r}^n, \mathbf{p}^n)}{\partial \mathbf{p}_i}; \quad \dot{\mathbf{p}} = -\frac{\partial H(\mathbf{r}^n, \mathbf{p}^n)}{\partial \mathbf{r}_i}. \tag{A.2}$$

In most molecular dynamics (MD) simulations, one assumes $\mathbf{p}_i = m_i \dot{\mathbf{r}}_i$, so that the system is governed by the laws of Newtonian mechanics. Several well-documented and verified MD

codes are in wide-spread use for such calculations [12, 65, 18, 31, 30, 52, 11] and generally provide access to interaction potentials of the form

$$V(\mathbf{r}^n) = \sum_{i=1}^{n} V_1(\mathbf{r}_i) + \sum_{\substack{i,j=1 \\ j<i}}^{n} V_2(\mathbf{r}_i, \mathbf{r}_j) + \sum_{\substack{i,j,k \\ k>j>i}}^{n} V_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \dots \tag{A.3}$$

where V_s is the s-body potential. Typically, cut-off radii are introduced to include only interactions of particles within that radius. A general form of V for a network of long molecular chains is [35]:

$$\begin{aligned} V(\mathbf{r}^n) \quad = \quad & \sum_{i=1}^{N_{co}} \frac{k_i}{2}(\mathbf{r} - \mathbf{r}_{0i})^2 + \sum_{i=1}^{N_\theta} \frac{\kappa_i}{2}(\theta_i - \theta_{0i})^2 \\ & + \sum_{i=1}^{N_\omega} \frac{\kappa_i^t}{2}(1 + \cos(i\omega - \gamma))^2 \\ & + \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left\{ 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right\} \end{aligned} \tag{A.4}$$

Here, $\mathbf{r}_{0i}$, θ_{0i}, and γ are initial values of $\mathbf{r}_i$, θ_i and ω_i, $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, N_{co} is the number of covalent bonds with bond stiffness k_i, κ_i denotes the rotational stiffness at the junction of bond lengths due to rotation θ_i of one link with respect to another, κ_i^t are the torsional stiffnesses, with torsional rotation ω, and the final term represents non-bonded interactions of a van der Waals bond represented by the 12-6 Lennard-Jones potential and the unbonded Coulombic interactions between atoms or molecules with charges q_i and q_j.

The molecular model parameters are thus the set $\boldsymbol{\mu}$:

$$\boldsymbol{\mu} = \{k_i, \kappa_i, \kappa_i^t, \varepsilon_{ij}, \gamma, \sigma_{ij}\} \tag{A.5}$$

and must be supplied to define $V(\mathbf{r}^n)$ for a given polymer realization. The initial positions $\mathbf{r}_{0i}$ of particle sites maybe regarded as known features of the model input independently of energy potentials, or they can be listed as additional parameters. We shall assume these quantities are exactly specified and write henceforth $V = V(\boldsymbol{\mu}; \mathbf{u}^n)$ to emphasize this point.

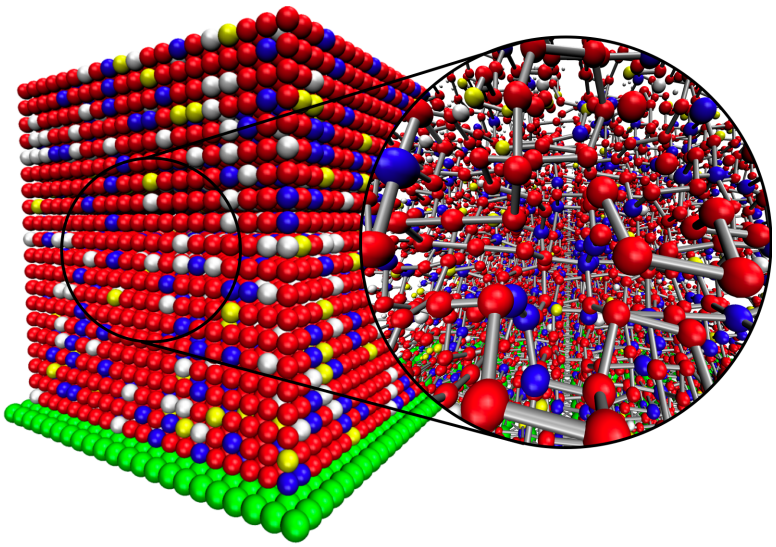


Figure A.1: A color-coded representation of one realization of a polymerization in which monomers, reactants, and voids are colored differently and assigned lattice sites in a KMC process, the green molecules representing the boundary substrate. Also shown is a blow up of bonds connecting sites once interaction potentials have been designated (Reproduced with permission from [9]).