Estimation and Control of Modeling Error: A General Approach to Multiscale Modeling

J.T. Oden, S. Prudhomme, P.T. Bauman, and L. Chamoin

Institute for Computational Engineering and Sciences
The University of Texas at Austin
1 University Station C0200, Austin, TX 78712, U.S.A.

CLARENDON PRESS • OXFORD

2008

CONTENTS

1

ESTIMATION AND CONTROL OF MODELING ERROR : A GENERAL APPROACH TO MULTISCALE MODELING

1.1 Problem setting

The following points provide the setting for the theory and methodologies described in this chapter.

- We wish to develop reliable and predictive computer simulations of the behavior of very large and complex molecular and atomistic systems. Such systems are encountered with increasing frequency in nanomanufacturing, in the design of advanced materials, in semiconductor manufacturing, in the analysis of biological systems, in drug design, and in numerous bio-medical applications. The simulation of the behavior of such systems is a critical challenge facing advancements in many areas of science and engineering.
- Detailed computational models of such large systems may involve hundreds of millions of unknowns. There may also be uncertainties in the parameters defining the models of such systems, so the unknowns are generally random variables and the models are stochastic. Such problems are well outside the capabilities of the largest and fastest computers that exist today or are likely to exist for many decades.
- In all computer simulations of physical systems, there are generally specific features of the solution that are of primary interest, the so-called *quantities of interest* (QoI's) or target outputs. It is generally assumed that in many cases these quantities are largely dependent on local fine-scale features of the model, and that at distances remote from those at which the quantities of interest are defined, behavior is determined by coarser scales representing averages in some sense of the fine-scale behavior. Thus, multiscale models are needed to reduce the size of the problem to one in which only the phenomena at specific scales affect the accuracy of the quantities of interest. This is really the only reason to consider multiscale modeling : to include in the computational model only the scales, and correspondingly the numbers of unknowns, needed to deliver quantities of interest with sufficient accuracy. But how does one know what level of fine-scale or coarse-scale information is needed in a model to obtain approximations of the QoI's with sufficient accuracy ?
- The only way to resolve this last question is to develop methods of estimating the error produced in quantities of interest by averaging or filtering out fine-scale efffects. Such *a posteriori* error estimates clearly pertain to the relative error between the full fine-scale (and generally intractable) model and

other approximations of coarser-scale. Coarse-scale models with dramatically fewer numbers of degrees of freedom compared to the base fine-scale model are obtained by various coarsening methods and the terms coarse graining, upscaling, homogenization, dimensional reduction, etc., are used to describe such processes. Many of such methods are *ad hoc* and do not attempt to estimate and control errors due to coarsening.

- We are interested here in a general class of methods for multiscale modeling that derive from the general setting just described :

  1. they assume the existence of a well-posed fine-scale base model of molecular systems that is generally intractable but of sufficient detail and sophistication to capture all events of interest with acceptable accuracy ;
  2. specific quantities of interest are identified that are representable as functionals of the fine-scale solution ;
  3. various averaging techniques are used to produce coarser-scale models in particular subdomains of the solutions of the fine-scale model, and, hence, hybrid models of multiple scales may be produced, and these models are tractable ;
  4. the coarse-grained hybrid models are "solved" and approximations of the quantities of interest are computed using the (incorrect) solutions of the hybrid models ;
  5. *a posteriori* estimates of error in the hybrid-coarse-scale QoI's compared to the actual fine-scale values are computed ; if they are small (compared to a preset tolerance level), the analysis is terminated and the hybrid model is accepted as a sufficiently accurate approximation of the fine-scale model ; if the error is large, the hybrid model contains insufficient fine-scale information and it must be refined to reduce the error in the QoI's ;
  6. the hybrid model is adapted by the addition of fine-scale features in appropriate subdomains in a way that systematically reduces the estimated error in the QoI's until error tolerances are met ; the process terminates when a near-optimal multiscale model is generated which yields acceptable values of the QoI's.

  These six steps characterize the *Goals Algorithm* for adaptive modeling. We discuss averaging (coarse-graining) techniques, error estimation, the adaptive strategy, and the all-important coupling algorithms that provide interfaces between the fine and coarse-scale regions of the model in subsequent sections.

We will focus on a specific class of molecular models : lattice-based models of polymer systems in static equilibrium (zero temperature). These types of molecular statics problems are encountered in nanomanufacturing of semiconductor devices using imprint lithography, an application addressed in the dissertation

of Bauman (2008) and in related publications (e.g. Bauman *et al.* 2008*a*, 2008*b*; Oden and Prudhomme 2002, Oden *et al.* 2006, Prudhomme *et al.* 2008*a*). The general methodology described is also applicable to problems of molecular dynamics, as shown in (Oden *et al.* 2006), but the polymer equilibrium problem has features of particular interest in multiscale modeling : the calculation of the structure and properties of the molecular system itself through the chemical process of polymerization. We also discuss this aspect of the modeling process.

1.2 The general theory of modeling error estimation

The idea of replacing a general mathematical model of physical events (the fine-scale *base model*) with a surrogate with coarser features and potentially fewer degrees of freedom (coarse graining, etc.) can be understood in a quite general abstract setting. We wish to find the vectors $\mathbf{u}$ in some topological vector space of trial functions U such that

$$A(\beta)\mathbf{u} = \mathbf{F} \tag{1.1}$$

where $A : U \longrightarrow V'$ is a nonlinear operator, possibly dependent on a set of parameters β, and $\mathbf{F}$ is data that belongs to topological dual V' of a space of test vectors V. The boundary and initial conditions pertinent to the model are embedded in the definitions of U and $\mathbf{F}$. We assume that solutions $\mathbf{u} = \mathbf{u}(\beta)$ exist for each $\mathbf{F}$ and for all β in some subset of parameter space in which (1.1) remains well-posed, but that the solutions may not be unique, and, of course, that they depend upon the data β. In general, the parameters are random variables and, therefore, $\mathbf{u}$ is a random variable. If it were possible to solve (1.1) for $\mathbf{u}(\beta)$ (which, in general, it is not), we are primarily interested in calculating particular quantities of interest (QoI's) that can be represented as functionals on U :

$$\boxed{\begin{array}{l}\text{Quantity of Interest} \quad Q(\mathbf{u}(\beta)) \\ Q : U \longrightarrow \mathbb{R} \\ (\mathbf{u}(\beta) \text{ is a solution of (1.1)})\end{array}} \tag{1.2}$$

For example, (1.1) may represent a large model of the dynamics of a complex molecular system, with $\mathbf{u}$ the set of trajectories of individual molecules over a specified time interval, and $Q(\mathbf{u})$ could represent the spatial and temporal average of the motion of a small subset of molecules over a time interval. The general problem (1.1) and (1.2), of course, can represent virtually any class of problems in mathematical physics, including problems in continuum mechanics.

Since (1.1) is generally unsolvable, we replace it with an approximate model, here perhaps one with coarser-scale features, but one that we presume can be solved in some sense :

$$A_0(\beta_0)\mathbf{u}_0 = \mathbf{F} \tag{1.3}$$

This model is a surrogate to (1.1), involving an operator $A_0 : U \longrightarrow V'$ which approximates A in some sense, and a set of parameters β_0. We hope that the solution $\mathbf{u}_0$ is "close" to $\mathbf{u}(\beta)$ in some sense. Obviously, this "closeness" makes sense

only in regard to quantities of interest. Thus, upon solving (1.3), we compute $Q(\mathbf{u}_0)$. The modeling error is then

$$\boxed{\varepsilon_0 = Q(\mathbf{u}(\beta)) - Q(\mathbf{u}_0(\beta_0))} \tag{1.4}$$

How well the surrogate model approximates the base model is determined by the magnitude of ε_0 and whether or not it is small enough for the application at hand. The error ε_0 is thus also a guide to coarse graining and dimensional reduction ; it quantifies exactly the effect of any particular choice of local averaging, coarse graining, etc., on values of the principal targets of the simulation : the QoI's.

The Error Estimate. Remarkably, if the data of the base problem are known (but not the solution $\mathbf{u}(\beta)$) the error ε_0 can be estimated provided one computes the solution of an adjoint problem corresponding to the particular quantity of interest Q. The basic steps are as follows :

- Since $A(\beta)\mathbf{u}-\mathbf{F}$ lies in the dual space V', we may view (1.1) as the condition that the residual functional $R : V \longrightarrow \mathbb{R}$ vanishes :

$$R(\beta, \mathbf{u}; \mathbf{v}) = \langle A(\beta)\mathbf{u} - \mathbf{F}, \mathbf{v}\rangle = 0 \quad \forall \mathbf{v} \in V \tag{1.5}$$

  Here $\langle\cdot,\cdot\rangle$ denotes duality pairing on $V'\times V$. We employ here the convention that functionals are linear in arguments to the right of the semi-colon, but possibly nonlinear functions of those entries to the left of the semi-colon. Hence, R is a linear functional of the test vector $\mathbf{v}$.
- The *adjoint problem* consists of finding $\mathbf{p} \in V$ such that

$$A'(\beta, \mathbf{u})^T\mathbf{p} = Q'(\beta, \mathbf{u}) \quad \text{in } V' \tag{1.6}$$

  where $A'(\beta, \mathbf{u})$ is the *linear* operator defined by

$$\lim_{\theta\to 0}\frac{1}{\theta}\langle A(\beta)(\mathbf{u} + \theta\mathbf{p}) - A(\beta)\mathbf{u} - \theta A'(\beta, \mathbf{u})\mathbf{p}, \mathbf{v}\rangle = 0 \quad \forall \mathbf{v} \in V$$

  and $Q'(\beta, \mathbf{u})$ is the linear functional,

$$\langle Q'(\beta, \mathbf{u}), \mathbf{v}\rangle = \lim_{\theta\to 0}\frac{1}{\theta}\big(Q(\beta, \mathbf{u} + \theta\mathbf{v}) - Q(\beta, \mathbf{u})\big)$$

  The solution $\mathbf{p}$ of the adjoint problem may be thought of as the *generalized Green function* corresponding to the particular choice Q of the QoI. It is also interpreted as the Lagrange multiplier associated with minimizing $Q(\mathbf{v})$ subject to the constraint (1.1).
- With $R(\beta, \mathbf{u}; \cdot)$ given by (1.5) and $\mathbf{p}$ the solution of (1.6) for given Q, it is shown in (Oden and Prudhomme 2002) that

$$\boxed{\varepsilon_0 = R(\beta, \mathbf{u}_0; \mathbf{p}) + \Delta} \tag{1.7}$$

  where Δ is a remainder depending on terms of quadratic and higher order in the error ε_0. Our assumption is that if the surrogate model (1.3) is close enough to (1.1), Δ is negligible and

$$\varepsilon_0 \approx R(\beta, \mathbf{u}_0; \mathbf{p}) \tag{1.8}$$

Our goal is to derive a family of algorithms that generate a sequence of surrogates and a sequence of approximations of the adjoint and the residual that reduces the modeling error ε_0 to a tolerable level.
We remark that the functionals in (1.5) and (1.6) define semilinear and linear forms,

$$\begin{aligned} B(\beta, \mathbf{u}; \mathbf{v}) &= \langle A(\beta)\mathbf{u}, \mathbf{v}\rangle \\ F(\mathbf{v}) &= \langle F, \mathbf{v}\rangle \\ B'(\beta, \mathbf{u}; \mathbf{v}, \mathbf{p}) &= \langle A'(\beta, \mathbf{u})^T\mathbf{p}, \mathbf{v}\rangle \\ Q'(\beta, \mathbf{u}; \mathbf{v}) &= \langle Q'(\mathbf{u}(\beta)), \mathbf{v}\rangle \end{aligned} \tag{1.9}$$

so that the primal and adjoint problems can be written as

$$\boxed{\begin{aligned} &\text{Find } (\mathbf{u}, \mathbf{p}) \in U \times U \\ &\qquad B(\beta, \mathbf{u}; \mathbf{v}) = F(\mathbf{v}) \qquad \forall \mathbf{v} \in V \\ &\qquad B'(\beta, \mathbf{u}; \mathbf{v}, \mathbf{p}) = Q'(\mathbf{u}(\beta), \mathbf{v}) \qquad \forall \mathbf{v} \in V \end{aligned}} \tag{1.10}$$

The rates of change (Gâteaux derivatives) of these forms due to variations in the parameters β can also be computed

$$\begin{aligned} B_\beta(\beta, \mathbf{u}; \mathbf{v}, w) &= \lim_{\theta \to 0} \theta^{-1}\big(B(\beta + \theta w, \mathbf{u}(\beta + \theta w); \mathbf{v}) - B(\beta, \mathbf{u}; \mathbf{v})\big) \\ Q_\beta(\beta, \mathbf{u}; \mathbf{v}, w) &= \lim_{\theta \to 0} \theta^{-1}\big(Q(\beta + \theta w, \mathbf{u}(\beta + \theta w); \mathbf{v}) - Q(\beta, \mathbf{u}; \mathbf{v})\big) \end{aligned}$$

1.3 A large-scale molecular statics model

While the framework described up to this point is quite general, we shall focus on a class of problems in nanomanufacturing that exhibit many of the features and challenges of multiscale modeling at the molecular level. The target application is Step and Flash Imprint Lithography, a process of stamping polymer etch barriers at room temperature to produce high-precision features of semiconductor devices. Full details of this process are discussed in (Bauman 2008). For our present purposes, it suffices to assert that the process involves creating a polymer through a chemical process that fills a quartz template designed to mold the surface features of devices with dimensions on the order of 40-70 nanometers or smaller. We wish to model the creation of the polymer and its densification (static deformation to an equilibrium configuration) using a lattice representation of the polymer molecules and chains, ignoring thermal effects. While such lattice models of polymer statics are often used (e.g. Vanderzande 1998), much more complex models could also be used to define the base model. Our adaptive modeling procedure also remains valid for such systems.

Figures 1.1 and 1.2 show the color-coded molecule locations in a lattice and the network of polymer chains arising from one realization of the polymerization process. For semiconductor units of dimensions $70 \times 200 \times 1000$ nanometers, these models can involve up to 10^7 unknown site displacement components.

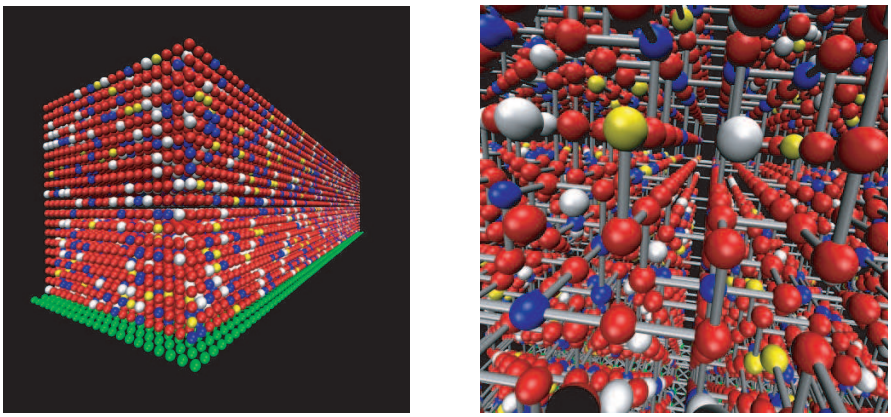

FIG. 1.1. Lattice model of one realization of the polymerization of a system of two monomers, with cross-links, reactants, and voids.

1.4 The family of six algorithms governing multiscale modeling of polymer densification

Six major algorithms are developed and implemented to model the molecular statics of lattice-type models of polymer systems. These are listed below in the order of their implementation.

1. *Polymerization : Kinetic Monte Carlo method.* A liquid solution of specific monomers, cross-link molecules, reactants, each with given initial volume fractions, is subjected to ultraviolet light, which initiates a chemical reaction. As a result, a network of polymer chains and cross-links is created, the chains representing long molecules with monomer links connected by covalent bonds. Other bonds, such as Van der Waals bonds, may also be created across chains. The process is called *polymerization.* Possible conformations of molecules and the configuration of the polymer structure are determined using a biased Monte Carlo algorithm in which molecular constituents are randomly distributed over a 3D lattice in proportion to the initial volume fraction. A reactant or molecule with free electrons existing in one cell is allowed to react with those of neighboring cells which have the highest probability of reaction as determined by the Arrhenius law,

$$R = Ce^{-kE_aT}$$

k being the reaction rate and E_a the activation energy. The probability of a reaction is proportional to the rate k :

$$p \propto k$$

and the rates are experimentally determined (Long *et al.* 2007). This aspect of the algorithm is similar to the classical Metropolis method (Metropolis

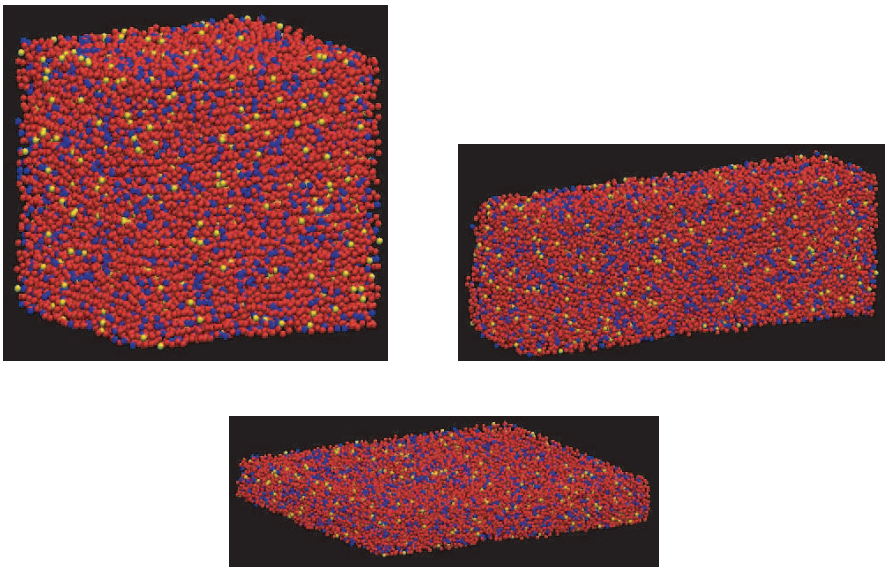

FIG. 1.2. Virtual experiments on a computer-generated Representative Volume Element (RVE) of a polymer : relaxation (top left), uniaxial stretch (top right), biaxial stretch (bottom).

et al. 1953). At the conclusion of n steps of this process, one realization of the polymer, such as that shown in Fig. 1.1, is obtained. Complete details of this algorithm are given in (Bauman 2008).

2. *Molecular potentials.* With a given molecular distribution over the polymer determined by Algorithm 1, molecular potentials are selected to represent the intermolecular forces. In this work, covalent bonds on polymer chains are represented by harmonic potentials,

$$V(\mathbf{r}) = \frac{k}{2}(\mathbf{r} - \mathbf{r}_0)^2$$

weak Van Der Waals bonds are represented by 6-12 Lennard-Jones potentials, etc. For simplicity, only nearest neighbors are accounted for here, but in principal, long-range effects could easily be included at the cost of greater computational complexity.

3. *Densification Algorithm : Inexact Newton-Raphson with trust region.* It is only at this stage of the analysis that we actually arrive at the possible base model of equilibrium configurations of the polymer structure, and this for only one realization of the polymer. For this realization, we are faced with the problem of solving very large nonlinear algebraic systems of the form (1.10) with

$$\left.\begin{aligned} B(\beta,\mathbf{u};\mathbf{v}) &= \sum_{i=1}^{N}\sum_{k=1}^{N_i}\frac{\partial E_{ik}(\beta;\mathbf{u})}{\partial \mathbf{u}_i}\cdot\mathbf{v}_i \\ F(\mathbf{v}) &= \sum_{i=1}^{N}\mathbf{f}_i\cdot\mathbf{v}_i \end{aligned}\right\} \tag{1.11}$$

where $\mathbf{u}$ is a vector of lattice site displacements $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N)$, $\mathbf{v}$ is a set of N test vectors, $E_{ik}(\beta;\mathbf{u})$ is the energy associated with site $\mathbf{x}_i$ and N_i neighbors $\mathbf{x}_k$, for a particular polymer realization, and $\mathbf{f}_i$ is the applied force at site $\mathbf{x}_i$. We solve the nonlinear system,

$$\sum_{k=1}^{N_i}\left(\frac{\partial E_{ik}(\beta;\mathbf{u})}{\partial \mathbf{u}_i} - \mathbf{f}_i\right) = 0 \quad 1 \le i \le N \tag{1.12}$$

by the Inexact Newton Trust Region method or we solve, for each β, the equivalent optimization problem,

$$\mathbf{u} = \arg\min_{\mathbf{v}} E(\mathbf{v}) \tag{1.13}$$

where $E(\mathbf{v}) = \sum_{i=1}^{N}\sum_{k=1}^{N_i} E_{ik}(\beta;\mathbf{v}) - \sum_{i=1}^{N}\mathbf{f}_i\cdot\mathbf{v}_i$ using the TAO/PETSc codes (Balay *et al.* 2001, Benson *et al.* 2007).

4. *Algorithms for computing surrogate models.* To reduce the enormous size of the systems of equations to a manageable level, we construct a coarse-scale continuum model of the polymer using virtual experiments on Representative Volume Elements (RVE's) and we construct an interface region between the discrete polymer domain and the continuum.

   (a) *Virtual experiments.* The polymerization routine is used to generate a sequence of realizations of a cubic RVE of the polymer. Owing to the choice of monomers, we know that rate effects are minimal, and can be neglected, and that the imprint process is performed at constant room temperature. No biases in the polymerization can lead to macroscopic heterogeneities or anisotropies over many realizations. Thus, we assume that at the macroscale level, the corresponding continuum is characterized as an isotropic, homogeneous, compressible hyperelastic material with a stored energy functional of the form,

$$\begin{aligned} W = \hat{W}(I_1, I_2, I_3) =& \alpha(I_1 - 3) + \beta(I_2 - 3) + \gamma(J-1)^2 \\ & - (2\alpha + 4\beta)\ln J \end{aligned} \tag{1.14}$$

   where α, β, γ are material constants, I_1, I_2 are the first two principal invariants of the deformation tensor $\mathbf{C} = \mathbf{F}^T\mathbf{F}$ and $J = \det\mathbf{F}$, $\mathbf{F}$ being the deformation gradient. For each realization, we subject the RVE to independent homogeneous deformation (stretch, lateral compression, etc.) to determine histograms of the material constants α,

β, γ as indicated in Fig. 1.2. We increase the dimensions of the RVE for each realization until essentially no changes in the experimentally-determined parameters occur.

(b) *The interface.* We construct an overlap domain that provides an interface between regions where the molecular model exists and the continuum model generated in step a) above. This is accomplished using the discrete-continuum Arlequin method (Ben Dhia and Rateau 2005).

(c) *Residual force calculation.* At the conclusion of the polymerization step, residual forces (and "strains") exist in the molecular model that force it to comply to the regular geometry shape of the lattice relative to a cartesian coordinate system. We account for these residual forces using an algorithm developed by Bauman (2008). We consider again a cubic RVE, compute the forces on the faces required to maintain the cubic shape, and apply equal-and-opposite forces to make the RVE forces free of external forces. The densification of the cube then takes it into a configuration with no forces on the molecules on the original RVE faces. Virtual experiments are implemented using these relaxed configurations. Details of this important process are given in (Bauman 2008).

5. *The adjoint problem.* We identify one or more quantities of interest (QoI's). For example, the slump of a patch Ω of a polymer device after it is removed from a template can be characterized by the functional

$$Q(\mathbf{u}) = \frac{1}{|\Omega|} \sum_{i \in I} \mathbf{u}_i \cdot \mathbf{k} \tag{1.15}$$

where $|\Omega|$ is the area of the surface, I an index set of lattice sites in Ω, and $\mathbf{k}$ a unit exterior vector normal to the exterior face of Ω. The adjoint problem corresponding to (1.15) is

$$B'(\beta, \mathbf{u}; \mathbf{v}, \mathbf{p}) = Q(\mathbf{v}) \quad \forall \mathbf{v} \in V \tag{1.16}$$

where now

$$B'(\beta, \mathbf{u}; \mathbf{v}, \mathbf{p}) = \sum_{i=1}^{N} \sum_{k=1}^{N_i} \frac{\partial^2 E_{ik}(\beta; \mathbf{u})}{\partial \mathbf{u}_i \partial \mathbf{u}_k} \mathbf{v}_i \cdot \mathbf{p}_k \tag{1.17}$$

This is the adjoint problem corresponding to the full fine-scale base problem (1.12). In general, it cannot be solved because the solution $\mathbf{u}$ to the primal problem is not known. Instead, we may solve a surrogate problem for an approximation $\hat{\mathbf{p}}$ of $\mathbf{p}$:

$$B'(\beta, \hat{\mathbf{u}}; \mathbf{v}, \hat{\mathbf{p}}) = Q(\mathbf{v}) \quad \forall \mathbf{v} \in V \tag{1.18}$$

Here $\hat{\mathbf{u}}$ is a solution of one of a sequence of approximations of $\mathbf{u}$ generated using the goals algorithm.

6. *The goals algorithm.* For simplicity, consider the two-dimensional polymer lattice domain D shown in Fig. 1.3 and suppose that the quantity of interest involves the motion of lattice sites confined to a subdomain Ω, as shown. We construct a larger open domain Θ_1 containing Ω in which all of the fine-scale degrees of freedom are contained. The interface region Γ_1 is created using the Arlequin method that connects the particle domain Θ_1 with the continuum model domain $D_1 = D - \Theta_1$. The deformation of the continuum domain D_1 is modeled using the finite element method so that in the hybrid model, the solution is discrete and can be compared with $\mathbf{u}$ in the trial space U. Let $\hat{\mathbf{u}}^{(1)}$ be the solution of this hybrid model. We compute the residual

$$R^{(1)} = R^{(1)}(\hat{\beta}, \hat{\mathbf{u}}^{(1)}, \mathbf{p}^{(1)}) \approx \varepsilon_0^{(1)} \tag{1.19}$$

where $\mathbf{p}^{(1)}$ is the solution of a surrogate adjoint problem to be described momentarily. The subregion Θ_1 is a member of a set of mutually disjoint sets forming a partition of D :

$$\bar{D} = \bigcup_{k=1}^{N} \bar{\Theta}_k \quad , \quad \Theta_k \cap \Theta_j = 0, k \neq j$$

The total residual can be represented as the sum

$$R^{(1)} = \sum_{k=1}^{N} R_k^{(1)}$$

where $R_k^{(1)}$ is the portion of the residual contributed by subdomain Θ_k. We choose $\alpha \in (0, 1)$ and invoke the condition, refine every subdomain Θ_k such that

$$R_k^{(1)} \geq \alpha R_{max}^{(1)} \quad ; \quad R_{max}^{(1)} = \max_k R_k^{(1)}$$

By "refine" we mean introducing all of the fine-scale unknowns (the molecular displacements) associated with that subdomain. The union of Θ_1 and all Θ_k satisfying $R_k^{(1)} \geq \alpha R_{max}^{(1)}$ creates a new hybrid model, with particles in $\Theta_1 \cup \Theta_k$, and with a new Arlequin interface Γ_2. A new residual $R^{(2)}$ is computed. This process is repeated until a preset tolerance $|R^{(s)}| < \gamma_{tol}$ is met, or the process is terminated after a preset number of iterates.

In computing the residuals, the adjoint vectors $\mathbf{p}^{(k)}$ must generally be computed by solving an adjoint problem on a finer-scale model than that used to calculate the hybrid solutions. More details on that important feature of these methods is given in (Bauman 2008).

1.5 Representative results

Verification of Error Estimator and *Adaptive Strategy.* We reproduce here results of (Bauman 2008) and (Bauman *et al.* 2008*b*) designed to test the effectivity of the *a posteriori* error estimate for a large molecular system and to test

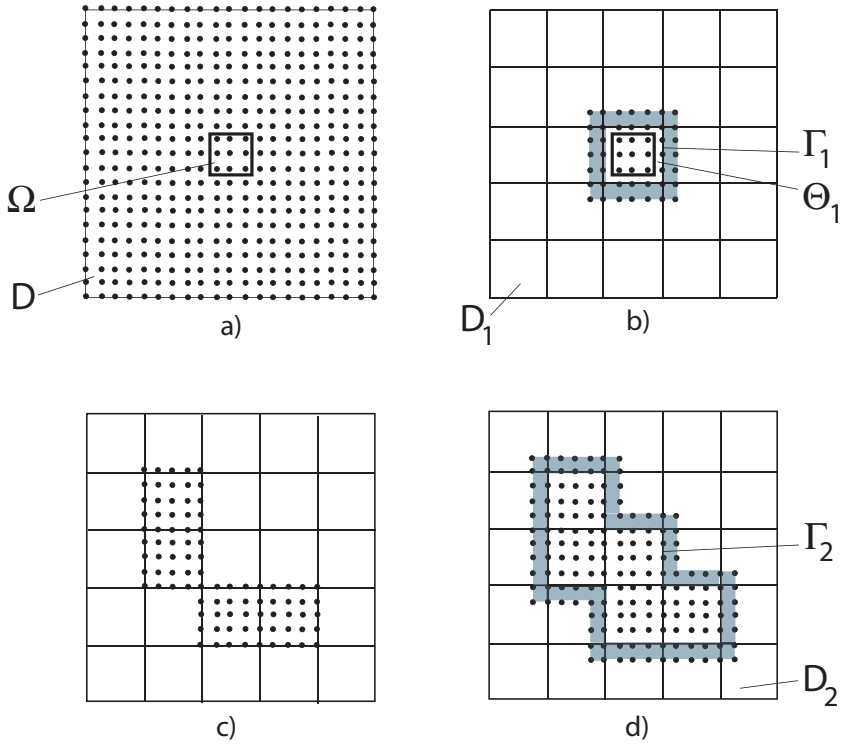

FIG. 1.3. The Goals Algorithm : a) a subdomain Ω associated with a quantity of interest in a fine-scale molecular model with domain D ; b) a hybrid model consisting of a partition of D with Θ_1 containing Ω, and interface region Γ_1 and a discretized continuum model D_1 ; c) subdomains with residual contributions exceeding or equal to αR_{max} ; and d) an adapted (second) hybrid model in the sequence of models.

the performance of the Goals Algorithm for a simplified case in which the solution of the base problem is actually known. As a model problem we consider a hypothetical case of a cubic $22 \times 22 \times 22$ lattice of molecules, fixed at its base, and undergoing a volume change, a 30% shrinkage during the densification of the polymer system. All covalent bonds are modeled using the harmonic potential with equal spring constants k. The resulting system has $31,944$ degrees of freedom. As a quantity of interest, we choose the slump of the polymer as represented by the average vertical displacement of a subset of molecules in a small region Ω centered at the top five lattice layers, as shown in Fig. 1.4. The average vertical displacement of the 5×5 patch of molecules at the center of the upper face of the cube will define the slump. A base solution $\mathbf{u}$ of the set of displacement vectors of the sites is computed using the Newton algorithm described earlier, and adjoint solution $\mathbf{p}$ corresponding to this particular quantity of interest is also computed.

The intensity of the adjoint solution vectors is shown color-coded in Fig. 1.5, the largest non-zero values in red near the quantity of interest subdomain Ω, and quickly dropping off to near zero at sites farther removed from the QoI.

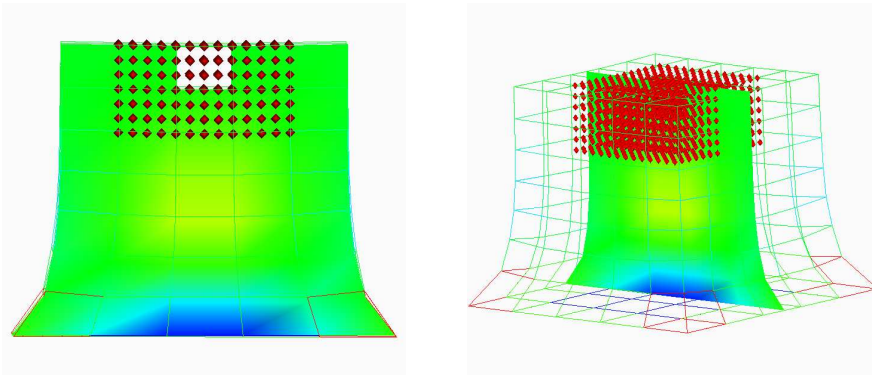

FIG. 1.4. The Arlequin problem.

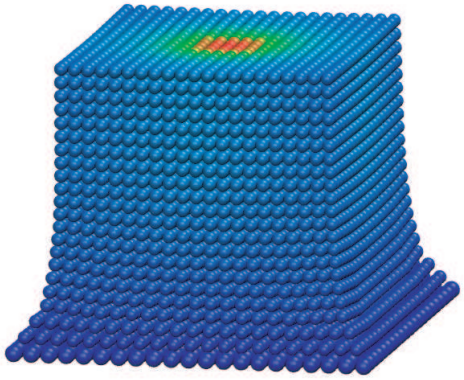

FIG. 1.5. Solution of the adjoint problem.

We next produce a sequence of surrogate models. The lattice domain is partitioned into 105 cubic subdomains, with 7 divisions in the vertical direction (normal to the fixed base) and 5×5 divisions in the planes perpendicular to the axis. These define the partition $\{\Theta_j\}$ used in the Goals Algorithm. As a first surrogate, we choose a subdomain, including Ω, of $13 \times 13 \times 7$ particles (molecules), outside of which the hyperelastic continuum, modeled as a compressible Mooney-Rivlin material described earlier, is approximated using only $5 \times 5 \times 7$ trilinear finite elements. A color-coded depiction of this partition is shown in Fig. 1.6, the pink denoting the fully discrete domain Ω, the red the overlap domain where the

Arlequin method imposes the connection of the fine-scale molecular model and the FEM-discretization of the hyperelastic continuum, which is green in the figure. The projection of the lattice site displacements within a continuum element to the current configuration of the surrogate lattice is denoted Π, as indicated in Fig. 1.7. Thus, if $\mathbf{u}_0$ is the solution of the discrete surrogate model, representing a set of displacement vectors at sites in the reduced model, $\Pi\mathbf{u}_0$ is the projection of this set of vectors on the lattice sites, enabling one to compare solutions of the surrogate to that of the base model.

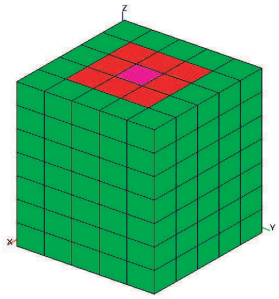

FIG. 1.6. Partition of the domain.

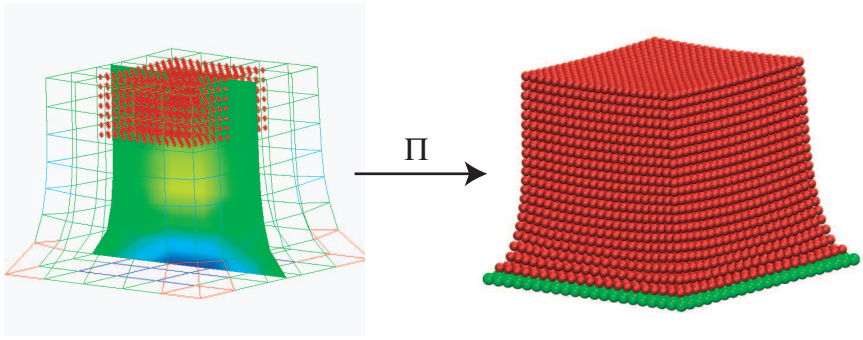

FIG. 1.7. Projection to the current configuration of the surrogate lattice.

A sequence of 5 model adaptations were used in the Goals Algorithm. Estimated errors were computed using the estimate,

$$\varepsilon_0 = R(\Pi\mathbf{u}_0, \beta; \mathbf{p})$$

with $\mathbf{p}$ the exact adjoint solution. The process reduced the error in the quantity of interest from near 10% to less than 5% in 5 steps (Tab. 1.1). The effectivity indices for the estimated error in each surrogate were very good, averaging over 90%. It is estimated that for this example, the remainder Δ in the estimate (1.7)

was around 8% of the total error. The sequence of surrogates generated by the Goals Algorithm with changing particle subdomains, Arlequin overlap domains, and discretized continuum elements, are shown in Fig. 1.8.

TABLE 1.1. Results obtained with the Goals Algorithm.

Adaptive Step	Error	Effectivity Index
0	9.77%	94.2%
1	8.24%	93.8%
2	7.71%	93.3%
3	6.75%	91.9%
4	5.08%	88.0%
5	4.74%	92.1%

There are many important details in implementing these types of adaptive modeling algorithms that can not be discussed in the limited space available here. These include such effects as refining the mesh approximating the continuum, varying the size of the Arlequin overlap domain, approximating the residual functional by evaluating it only over a subset of the partition domains, various approximations of the adjoint solution, and issues with developing scalable parallel code. These issues are discussed more fully in (Bauman 2008; Bauman *et al.* 2008*b*; Prudhomme *et al.* 2008*b*).

1.6 Extensions

We close this chapter with brief discussions of extensions of the method to important classes of problems in multiscale modeling. One area is the extension of goals algorithms to stochastic systems and the quantification of uncertainties in quantities of interest. The idea of quantifying uncertainty in errors in quantities of interest is discussed in (Oden *et al.* 2005). The difficult problem of extending the Arlequin-type methods for modeling the interface of discrete and continuum models to problems with random parameters and molecular structure is discussed in connection with model 1D problems in (Chamoin *et al.* 2008).

The framework developed earlier in this chapter can be extended to time-dependent problems in molecular dynamics and, as an important step in producing predictive simulations, to problems of calibration, inverse analysis, and optimal control. The mathematical formulation follows again arguments typical in optimal control theory (e.g. Le Dimet and Shutyaev 2005). One such extension involves introducing the following generalizations of (1.9) and (1.10) :

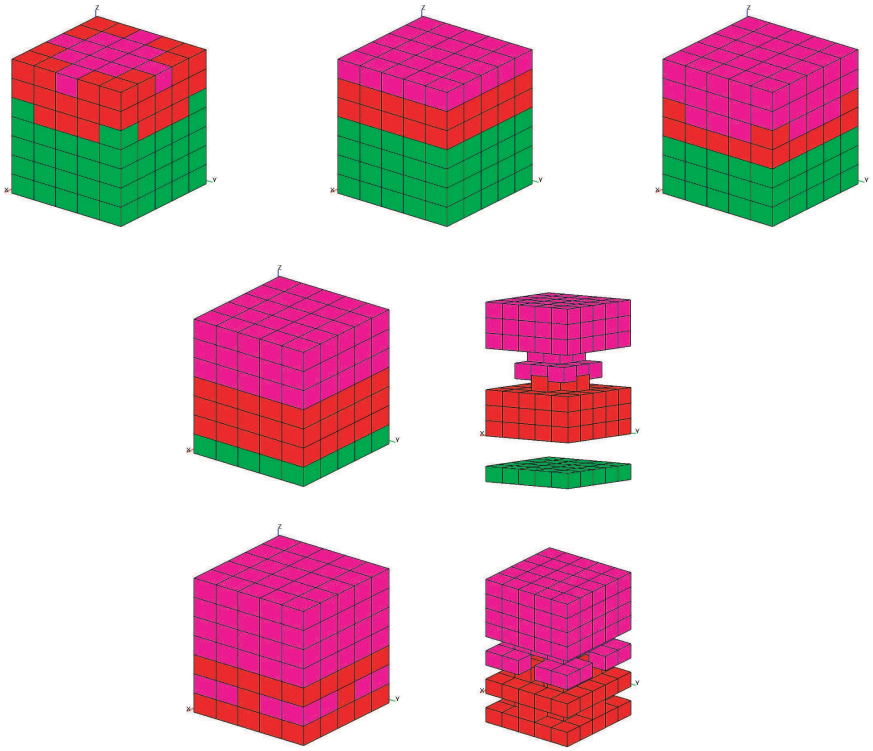

FIG. 1.8. Sequence of surrogates generated by the Goals Algorithm.

$$\begin{aligned} B(\beta,\mathbf{u};\mathbf{v}) &= \textstyle\int_0^T \mathbf{v}\cdot\big(M\ddot{\mathbf{u}} - A(\beta,\mathbf{u})\big)\mathrm{d}t + \mathbf{v}(0)\cdot M\dot{\mathbf{u}}(0) - \dot{\mathbf{v}}(0)\cdot M\mathbf{u}(0) \\ F(\beta,\mathbf{v}) &= \mathbf{v}(0)\cdot M\psi_0 - \dot{\mathbf{v}}(0)\cdot M\varphi_0 \\ B'(\beta,\mathbf{u};\mathbf{v},\mathbf{p}) &= \textstyle\int_0^T \mathbf{v}\cdot\big(M\ddot{\mathbf{p}} - A'(\beta,\mathbf{u})^T\mathbf{p}\big)\mathrm{d}t + \dot{\mathbf{v}}(T)\cdot M\mathbf{p}(T) - \mathbf{v}(T)\cdot M\dot{\mathbf{p}}(T) \\ \Gamma(\beta,\mathbf{v}) &= Q(\beta,\mathbf{v}) + \frac{\mu}{2}\|\varphi_0-\varphi\|_{\mathcal{X}}^2 + \frac{\lambda}{2}\|\psi_0-\psi\|_{\mathcal{Y}}^2 + \frac{1}{2}\|C\mathbf{v}-\gamma_{\mathrm{obs}}\|_{\mathcal{Z}}^2 \end{aligned} \tag{1.20}$$

Here M is the symmetric mass matrix for the entire molecular system and superimposed dots ($\dot{}$), ($\ddot{}$) indicate time derivatives. Thus, for instance,

$$\mathbf{v}\cdot M\ddot{\mathbf{u}} = \sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{v}_i\cdot M_{ij}\frac{\mathrm{d}^2\mathbf{u}_j}{\mathrm{d}t^2}$$

The initial velocity and displacement vectors are ψ_0 and φ_0 and are typically unknown. In the expression for Γ in (1.20), μ and λ are real positive parameters, and $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ are Hilbert spaces of initial displacement fields, velocity fields, and time-dependent vectors γ_{obs} of values of experimental observations of the system. Thus, $C : U \to \mathcal{Z}$ is a *calibration operator*, mapping vectors in the trial

space U into the observation space $\mathcal{Z}$. The functional $\Gamma : \mathbb{P} \times U \to \mathbb{R}$ is the cost functional providing control of the quantity of interest Q, itself a functional on U, the initial data, and the calibration data. The data vector $\beta \in \mathbb{R}^m$ of model parameters, primarily the parameters appearing in the molecular energy potentials and initial data, belongs to a manifold $\mathbb{P}$ of real, constant-in-time parameters, and discrete-valued functions defined on the initial configurations of the system. The mass matrix M_{ij} also may be random as different monomers, and hence different molecular masses can occupy site i for different realizations of the polymer structure. Thus β can be written,

$$\beta = (\omega, \varphi_0, \psi_0)$$

where ω is a vector of model data parameters, such as those appearing in the molecular potentials, and φ and ψ are the initial displacement and velocity vectors.

We are concerned with the optimal control problem,

$$\boxed{\begin{array}{l}\text{Find } (\beta, \mathbf{u}) \in \mathbb{P} \times U \text{ such that} \\ \Gamma(\beta, \mathbf{u}) = \inf_{\mathcal{W}} \Gamma(\gamma, \mathbf{v}) \\ \mathcal{W} = \{(\gamma, \mathbf{v}) \in \mathbb{P} \times U : B(\gamma, \mathbf{v}; \mathbf{w}) = F(\gamma; \mathbf{w}) \quad \forall \mathbf{w} \in V\}\end{array}} \tag{1.21}$$

The solution to (1.21) is characterized by the following system :

- The forward-primal problem :

$$\begin{aligned} \mathbf{u} &= \mathbf{u}(\beta) \in U \\ B(\beta, \mathbf{u}; \mathbf{v}) &= F(\beta; \mathbf{v}) \quad \forall \mathbf{v} \in V \end{aligned} \tag{1.22}$$

- The backward or adjoint problem :

$$\begin{aligned} \mathbf{p} =& \mathbf{p}(\beta) \in U \\ B'(\beta, \mathbf{u}; \mathbf{v}, \mathbf{p}) =& Q'(\beta, \mathbf{u}; \mathbf{v}) + (C\mathbf{u} - \gamma_{\text{obs}}, C\mathbf{v})_{\mathcal{Z}} \quad \forall \mathbf{v} \in V \end{aligned} \tag{1.23}$$

- The sensitivity of the cost functional

$$\Gamma_\beta(\beta, \mathbf{u}, \delta\beta) + F_\beta(\beta, \mathbf{p}, \delta\beta) - B_\beta(\beta, \mathbf{u}; \mathbf{p}, \delta\beta) = 0 \quad \forall \delta\beta \in \mathbb{P} \tag{1.24}$$

The system (1.22)–(1.24) is equivalent to the dynamical problem,

$$\begin{aligned} M\frac{\mathrm{d}^2\mathbf{u}}{\mathrm{d}t^2} &= A(\beta, \mathbf{u}) \\ \mathbf{u}(0) = \varphi \quad &; \quad \dot{\mathbf{u}}(0) = \psi \end{aligned} \tag{1.25}$$

$$\begin{aligned} M\frac{\mathrm{d}^2\mathbf{p}}{\mathrm{d}t^2} &= A'(\beta, \mathbf{u})^T\mathbf{p}(t) + \mathbf{q}'(\beta, \mathbf{u}) + C^T(C\mathbf{u} - \gamma_{\text{obs}}) \\ \mathbf{p}(T) = 0 \quad &; \quad \dot{\mathbf{p}}(T) = 0 \end{aligned} \tag{1.26}$$

$$\begin{aligned} -B_\omega(\beta,\mathbf{u};\mathbf{p}) + Q_\omega(\beta,\mathbf{u}) &= 0 \\ -B_{\varphi_0}(\beta,\mathbf{u};\mathbf{p}) + Q_{\varphi_0}(\beta,\mathbf{u}) + \mu(\varphi_0-\varphi) - M\dot{\mathbf{p}}(0) &= 0 \\ -B_{\psi_0}(\beta,\mathbf{u};\mathbf{p}) + Q_{\psi_0}(\beta,\mathbf{u}) + \lambda(\psi_0-\psi) + M\mathbf{p}(0) &= 0 \end{aligned} \tag{1.27}$$

where now

$$\begin{aligned} Q'(\beta,\mathbf{u};\mathbf{v}) &= \int_0^T \mathbf{q}'(\beta,\mathbf{u})\cdot\mathbf{v}\,\mathrm{d}t \\ A(\beta,\mathbf{u}) &= \sum_{i=1}^{N}\sum_{j=1}^{N_i}\frac{\partial E_{ij}(\beta,\mathbf{u})}{\partial \mathbf{u}_i} \\ A'(\beta,\mathbf{u})^T\mathbf{p} &= \sum_{i=1}^{N}\sum_{j=1}^{N_i}\sum_{k=1}^{N}\frac{\partial^2 E_{ij}(\beta,\mathbf{u})}{\partial \mathbf{u}_k\partial \mathbf{u}_i}\mathbf{p}_k \end{aligned} \tag{1.28}$$

To compute sensitivities of the solution $\mathbf{u}$ to changes in the parameters β, we consider a path $S(\beta)$ in U, parameterized by β, along which the forward problem is solved; i.e. $\forall\beta\in S(\beta)$, we have $B(\beta,\mathbf{u};\mathbf{v}) = F(\beta,\mathbf{v}) \quad \forall\mathbf{v}\in V$. The change in $\mathbf{u}$ due to the change in β is denoted $\mathbf{u}'(\beta)$. Thus, along $S(\beta)$, the change in $B(\beta,\mathbf{u};\mathbf{v}) - F(\beta,\mathbf{v})$ is zero :

$$B_\beta(\beta,\mathbf{u};\mathbf{v},\delta\beta) + B'(\beta,\mathbf{u};\mathbf{u}',\mathbf{v}) - F_\beta(\beta,\mathbf{v},\delta\beta) = 0 \quad \forall\mathbf{v}\in V$$

The backward (adjoint) problem along $S(\beta)$ satisfies

$$B'(\beta,\mathbf{u};\mathbf{u}',\mathbf{p}) = \Gamma'(\beta,\mathbf{u};\mathbf{u}')$$

where $\Gamma'(\beta,\mathbf{u};\mathbf{u}')$ is given by the right-hand side of (1.23) with $\mathbf{v}=\mathbf{u}'$. Thus, the change in the objective functional Γ along $S(\beta)$ is

$$D_\beta\Gamma = \partial_\beta\Gamma_{|(\mathbf{u},\varphi,\psi)} - B_\beta(\beta,\mathbf{u};\mathbf{p},\delta\beta) + F_\beta(\beta,\mathbf{p},\delta\beta)$$

These functions define the sensitivity of the solution to the parameters β.

Let $(\mathbf{u}_0,\mathbf{p}_0)$ denote a solution pair to any surrogate model approximating the system (1.22)–(1.24) with a parameter set β_0. Then arguments similar to those leading to (1.8) lead to the a posteriori error estimate,

$$\Gamma(\beta,\mathbf{u}) - \Gamma(\beta_0,\mathbf{u}_0) \approx \mathcal{R}(\beta_0,\mathbf{u}_0;\mathbf{p}) \tag{1.29}$$

where $\mathcal{R}(\beta_0,\mathbf{u}_0;\mathbf{p})$ is the time-dependent residual :

$$\mathcal{R}(\beta_0,\mathbf{u}_0;\mathbf{p}) = F(\beta_0;\mathbf{p}) - B(\beta_0,\mathbf{u}_0;\mathbf{p}) + F_\beta(\beta_0;\mathbf{p},\beta-\beta_0) - B_\beta(\beta_0,\mathbf{u}_0;\mathbf{p},\beta-\beta_0) \tag{1.30}$$

The difference $\beta-\beta_0$ can, in principle, be estimated using the inverse analysis and calibration process alluded to earlier. Further refinements of the estimate (1.29) are believed to be possible, but await further research. This modeling error estimate can be readily used as a basis for adaptive modeling via the goals algorithm described earlier.

1.7 Concluding comments

An important goal of contemporary computational science is quantifiable predictability, the systematic prediction of the behavior of physical systems with quantifiable metrics of confidence and uncertainty. What is remarkable is that the core of the success of predictive methods is the selection and ultimately the control of the models used as the bases of simulation. It is hoped that the methodologies described here and their applications to large scale problems provide some bases for advancing this important subject.

The idea of estimating and controlling errors induced in computational models by involking assumptions is an extremely powerful concept and can bring a new level of sophistication and rigor to the analyses of the most complex problems in multiscale modeling. Many issues remain to be resolved in applications of these ideas to large-scale problems, and these will involve advances in goals-type algorithms used to implement adaptive modeling strategies. Extensions to stochastic models and the use of frameworks similar to those discussed in Section 1.6 will make feasible the development of methods for uncertainty quantifications in large multiscale problems.

Acknowledgements : The support of this work by the Department of Energy under contract DE-FG02-05ER25701 is gratefully acknowledged.

References

Balay, S., Buschelman, K., Gropp, W.D., Kaushik, D., Knepley, M.G., McInnes, L.C., Smith, B.F. and Zhang, H. (2001).
PETSc Web page : *http ://www.mcs.anl.gov/petsc.*

Bauman, P.T. (2008). *Adaptive Multiscale Modeling of Polymeric Materials Using Goal-Oriented Error Estimation, Arlequin Coupling, and Goals Algorithms.* PhD dissertation, The University of Texas at Austin.

Bauman, P.T., Ben Dhia, H., Elkhodja, N., Oden, J.T. and Prudhomme, S. (2008). On the application of the Arlequin method to the coupling of particle and continuum models. *Computational Mechanics* (to appear).

Bauman, P.T., Oden, J.T. and Prudhomme, S. (2008). Adaptive multiscale modeling of polymeric materials : Arlequin coupling and goals algorithm. *Computer Methods in Applied Mechanics and Engineering* (submitted).

Ben Dhia, H. and Rateau, G. (2005). The Arlequin method as a flexible engineering design tool. *International Journal for Numerical Methods in Engineering* **62**(11) :1442–1462.

Benson, S., McInnes, L.C., Moré, J., Munson, T. and Sarich, J. (2007). *TAO User Manual (Revision 1.9).* Mathematics and Computer Science Division, Argonne National Laboratory, ANL/MCS-TM-242.
Web page :*http ://www.mcs.anl.gov/tao.*

Chamoin, L., Oden, J.T. and Prudhomme, S. (2008). A stochastic coupling method for atomic-to-continuum Monte-Carlo simulations. *Computer Methods in Applied Mechanics and Engineering* (in press).

Le Dimet, F.X. and Shutyaev, V. (2005). On deterministic error analysis in variational data assimilation. *Nonlinear Processes in Geophysics*; **12** :481–490.

Long, B.K., Keitz, B.K. and Willson, C.G. (2007). Materials for step and flash imprint lithography (S-FIL). *Journal of Materials Chemistry* (in press).

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, M. and Teller, E. (1953). Equation of state calculations by very fast computing machines. *Journal of Chemical Physics*; **21** :1087.

Oden, J.T. and Prudhomme, S. (2002). Estimation of modeling error in computational mechanics. *Journal of Computational Physics*; **182** :496–515.

Oden, J.T., Babuska, I., Nobile, F., Feng, Y. and Tempone, R. (2005). Theory and methodology for estimation and control of errors due to modeling, approximation,and uncertainty. *Computer Methods in Applied Mechanics and Engineering*; **194**(42-44) :4506–4527.

Oden, J.T., Prudhomme, S., Romkes, A. and Bauman, P.T. (2006). Multiscale modeling of physical phenomena : adaptive control of models. *SIAM Journal of Scientific Computing*; **28**(6) :2359–2389.

Prudhomme, S., Ben Dhia, H., Bauman, P.T., Elkhodja, N. and Oden, J.T. (2008). Computational analysis of modeling error for the coupling of particle and continuum models by the Arlequin method. *Computer Methods in Applied Mechanics and Engineering* (to appear).

Prudhomme, S., Chamoin, L., Ben Dhia, H. and Bauman, P.T. (2008). An adaptive strategy for the control of modeling error in two-dimensional atomic-to-continuum coupling simulations (submitted).

Vanderzande, C. (1998). *Lattice models of polymers*. Cambridge University Press.