Predictability of Coarse-Grained Models of Atomistic Systems in the Presence of Uncertainty

J. Tinsley Oden

Institute for Computational Engineering and Sciences
The University of Texas at Austin

Belytschko Lecture

Theoretical and Applied Mechanics Program
Departments of Mechanical Engineering
and Civil and Environmental Engineering

Northwestern University
October 14, 2014

Acknowledgements:

Collaborators: Kathryn Farrell
Danial Faghihi

DOE-DE-SC009286 MMICC
AFOSR: FA9550-12-1-0484
DOE Eureka: DE-SC0010518

Outline

1. The Logic of Predictive Science: What is Predictive Science?

Predictive Science: the scientific discipline concerned with assessing the predictability of mathematical and computational models of physical events. It embraces the processes of model selection, calibration, validation, verification, and their use in forecasting features of physical events with quantified uncertainty.

Comprehensive Nuclear Test Ban: UN 1996
Space Shuttle Accident, 2003
Climate/Weather Prediction
Predictive Medicine
⋮

1. The Logic of Predictive Science: What is Predictive Science?

> Mathematical constructions based on physical principles or empirical relations-generally based on inductive theories which attempt to characterize abstractions of physical reality

Predictive Science: the scientific discipline concerned with assessing the predictability of mathematical and computational models of physical events. It embraces the processes of model selection, calibration, validation, verification, and their use in forecasting features of physical events with quantified uncertainty.

Comprehensive Nuclear Test Ban: UN 1996
Space Shuttle Accident, 2003
Climate/Weather Prediction
Predictive Medicine
⋮

1. The Logic of Predictive Science: What is Predictive Science?

The process of adjusting the parameters of a model to improve the agreement of model predictions with experimental measurements

Mathematical constructions based on physical principles or empirical relations-generally based on inductive theories which attempt to characterize abstractions of physical reality

Predictive Science: the scientific discipline concerned with assessing the predictability of mathematical and computational models of physical events. It embraces the processes of model selection, calibration, validation, verification, and their use in forecasting features of physical events with quantified uncertainty.

Comprehensive Nuclear Test Ban: UN 1996
Space Shuttle Accident, 2003
Climate/Weather Prediction
Predictive Medicine
⋮

1. The Logic of Predictive Science: What is Predictive Science?

The process of adjusting the parameters of a model to improve the agreement of model predictions with experimental measurements

Mathematical constructions based on physical principles or empirical relations-generally based on inductive theories which attempt to characterize abstractions of physical reality

The process of determining the accuracy with which a model can predict features of physical reality

Predictive Science: the scientific discipline concerned with assessing the predictability of mathematical and computational models of physical events. It embraces the processes of model selection, calibration, validation, verification, and their use in forecasting features of physical events with quantified uncertainty.

Comprehensive Nuclear Test Ban: UN 1996
Space Shuttle Accident, 2003
Climate/Weather Prediction
Predictive Medicine
⋮

1. The Logic of Predictive Science: What is Predictive Science?

The process of adjusting the parameters of a model to improve the agreement of model predictions with experimental measurements

Mathematical constructions based on physical principles or empirical relations-generally based on inductive theories which attempt to characterize abstractions of physical reality

The process of determining the accuracy with which a model can predict features of physical reality

Predictive Science: the scientific discipline concerned with assessing the predictability of mathematical and computational models of physical events. It embraces the processes of model selection, calibration, validation, verification, and their use in forecasting features of physical events with quantified uncertainty.

The Quantities of Interest (QoIs): the goals of the simulation

Comprehensive Nuclear Test Ban: UN 1996
Space Shuttle Accident, 2003
Climate/Weather Prediction
Predictive Medicine
⋮

1. The Logic of Predictive Science: What is Predictive Science?

The process of adjusting the parameters of a model to improve the agreement of model predictions with experimental measurements

Mathematical constructions based on physical principles or empirical relations-generally based on inductive theories which attempt to characterize abstractions of physical reality

The process of determining the accuracy with which a model can predict features of physical reality

Predictive Science: the scientific discipline concerned with assessing the predictability of mathematical and computational models of physical events. It embraces the processes of model selection, calibration, validation, verification, and their use in forecasting features of physical events with quantified uncertainty.

The Quantities of Interest (QoIs): the goals of the simulation

UQ: quantifying the uncertainty in predicted QoIs

Comprehensive Nuclear Test Ban: UN 1996
Space Shuttle Accident, 2003
Climate/Weather Prediction
Predictive Medicine
⋮

1. The Logic of Predictive Science: What is Predictive Science?

The process of adjusting the parameters of a model to improve the agreement of model predictions with experimental measurements

Mathematical constructions based on physical principles or empirical relations-generally based on inductive theories which attempt to characterize abstractions of physical reality

The process of determining the accuracy with which a model can predict features of physical reality

Predictive Science: the scientific discipline concerned with assessing the predictability of mathematical and computational models of physical events. It embraces the processes of model selection, calibration, validation, verification, and their use in forecasting features of physical events with quantified uncertainty.

The Quantities of Interest (QoIs): the goals of the simulation

UQ: quantifying the uncertainty in predicted QoIs

Comprehensive Nuclear Test Ban: UN 1996
Space Shuttle Accident, 2003
Climate/Weather Prediction
Predictive Medicine
⋮

Predictability requires **knowledge** of the physical laws that are proposed to explain realities and requires recognizing and quantifying uncertainties

The Nature of Science

Science - The activity concerned with the systematic acquisition of knowledge

The Pillars of Science - I. Theory - inductive hypotheses
II. Observation - experiments
III. Computational Science

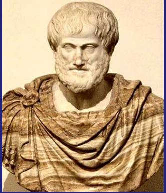
The Scientific Method - Test statements that are logical consequences of scientific hypotheses (theories) or related computer models and simulation through repeatable experiments or observations

Logic: The science dealing with the formal principles of reasoning (or the study of reasoning)

Deductive Reasoning (or deductive logic)

The process of reasoning from one or more general statements (axioms or premises) to reach logically **certain** conclusions

- "Top-down logic": premises $\Rightarrow$ conclusions
- Established as a formal discipline by Aristotle 384-322 B.C.

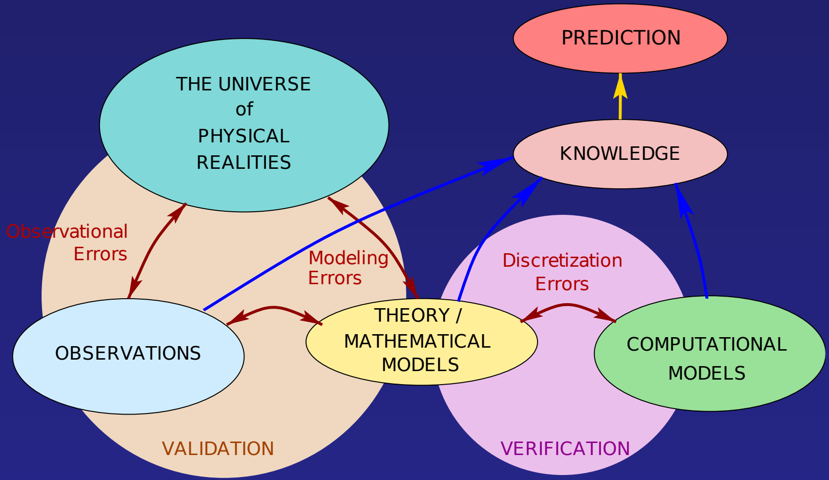


Inductive Reasoning (or inductive logic)

The process of reasoning by generalizing or **extrapolating** from initial information or **hypotheses**

- "Bottom-up logic": an open system including domains of epistemic *uncertainty* (allowing a conclusion to be false)

The Imperfect Paths to Knowledge



Cox's Theorem

Every natural extension of Aristotelian logic with uncertainties is Bayesian

Precisely:

There exists a continuous, strictly increasing, real-valued, non-negative function p, the plausibility of a proposition conditioned on information X, such that for every proposition A and B and consistent X

1. $p(A|X) = 0$ iff A is false given the information in X
2. $p(A|X) = 1$ iff A is true given the information in X
3. $0 \leq p(A|X) \leq 1$
4. $p(A \wedge B|X) = p(A|X)p(B|AX)$
5. $p(\bar{A}|X) = 1 - p(A|X)$ if X is consistent

Richard Trelked Cox, Am. J. Physics, 1946
Edwin T. Jaynes, Probability Theory: The Logic of Science, 2003
Kevin van Horn, J. Approx. Reasoning, 2003

Cox's Theorem

Every natural extension of Aristotelian logic with uncertainties is Bayesian

Precisely:

$\nexists\, A$ s.t. $(A|X)$ and $(\bar{A}|X)$ are true

There exists a continuous, strictly increasing, real-valued, non-negative function p, the plausibility of a proposition conditioned on information X, such that for every proposition A and B and consistent X

1. $p(A|X) = 0$ iff A is false given the information in X
2. $p(A|X) = 1$ iff A is true given the information in X
3. $0 \leq p(A|X) \leq 1$
4. $p(A \wedge B|X) = p(A|X)p(B|AX)$
5. $p(\bar{A}|X) = 1 - p(A|X)$ if X is consistent

Richard Trelked Cox, Am. J. Physics, 1946
Edwin T. Jaynes, Probability Theory: The Logic of Science, 2003
Kevin van Horn, J. Approx. Reasoning, 2003

Cox's Theorem

Every natural extension of Aristotelian logic with uncertainties is Bayesian

Precisely:

$\nexists\, A$ s.t. $(A|X)$ and $(\bar{A}|X)$ are true

There exists a continuous, strictly increasing, real-valued, non-negative function p, the plausibility of a proposition conditioned on information X, such that for every proposition A and B and consistent X

1. $p(A|X) = 0$ iff A is false given the information in X
2. $p(A|X) = 1$ iff A is true given the information in X
3. $0 \leq p(A|X) \leq 1$
4. $p(A \wedge B|X) = p(A|X)p(B|AX)$ — Bayes' Rule
5. $p(\bar{A}|X) = 1 - p(A|X)$ if X is consistent

Richard Trelked Cox, Am. J. Physics, 1946
Edwin T. Jaynes, Probability Theory: The Logic of Science, 2003
Kevin van Horn, J. Approx. Reasoning, 2003

Post-Cox Developments

- **Halpern, Joseph Y.**, Counterexample to Cox's Theorem - then a correction in an "Addendum to Cox's Theorem" (1999), then refuted by van Horn (2003)
- **Amborg, Stephan and Sjodin, Gunnar** (1999, 2000)
- **Van Horn, Kevin S.**, "Constructing a Logic of Plausibility - A Guide to Cox's Theorem," J. Approx. Reasoning (2003)
- **Jaynes, Edwin T.**, *Probability Theory: The Logic of Science* (2003)
- **Dupre, Maurice J. and Tipler, Frank J.**, "A Trivial Proof of Cox's Theorem" (2009)
- **McGrayne, Sharon B.**, *The Theory That Would Not Die* (2012)
- **Freedman, David** (1999, 2006)
- **Kleijn, B. J. K. and van der Vaart, A. W.**, *The Bernstein-von-Mises Theorem Under Misspecification* (2012)
- **Owhadi, Houman, Scoval, Clint and Sullivan, Tim**, "Bayesian Brittleness: Why no Bayesian Model is Good Enough" (2013)

The Logic of Science: Bayesian Inference

Bayes' Rule

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Thomas Bayes (1763):

"An Essay Towards Solving a Problem in the Doctrine of Chances" PRS

* Logical Probability $\supset$ frequency based theory

Bayesian Model Calibration, Validation, and Prediction

$\mathcal{P}(\Theta)$ = a parametric model class

$= \{A(\boldsymbol{\theta}, S, u(\boldsymbol{\theta}, S)) = 0\}$

mathematical model — parameters — scenario — solution

$$\pi(\boldsymbol{\theta}|\mathbf{y}) = \frac{\pi(\mathbf{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{y})}$$

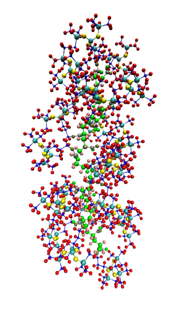
likelihood: $\pi(\mathbf{y}|\boldsymbol{\theta})$; prior: $\pi(\boldsymbol{\theta})$; posterior: $\pi(\boldsymbol{\theta}|\mathbf{y})$; evidence: $\pi(\mathbf{y})$

- Ω_i = reality

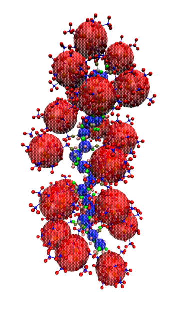
$$\Omega_i + \varepsilon_i = y_i$$
$$\Omega_i - d_i(u(\boldsymbol{\theta}, S)) = \eta_i$$

$$p(\varepsilon_i + \text{“}\eta_i\text{”}) = p(y_i - d_i(\boldsymbol{\theta})) = \pi(y_i|\boldsymbol{\theta}) = \text{likelihood}$$

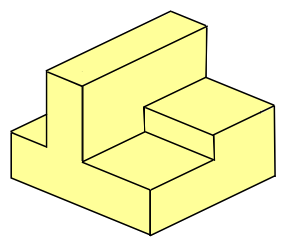
2. The Tyranny of Scales: Predictivity of Multiscale Models



⟹



⟹



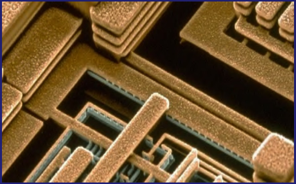
All-Atom (AA) Model

Coarse-Grained (CG) Model

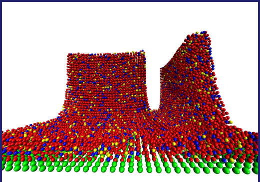
Macro (Continuum) Model

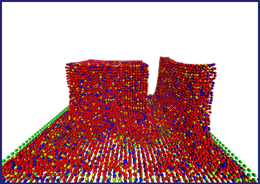
The confluence of all challenges in Predictive Science: Exactly what *is* the model? Is it "valid"? What is the level of uncertainty in the prediction?

Nanomanufacturing

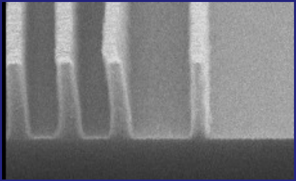


a) Semiconductor Component





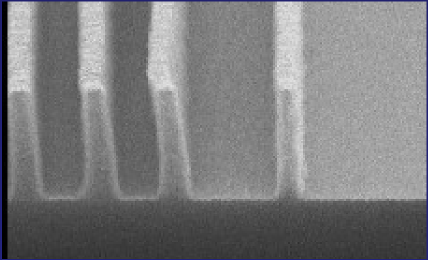
b) Multiblock Component

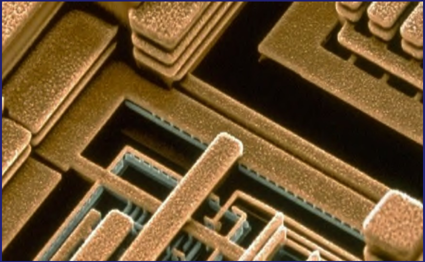


c) Manufacturing detail

National Medal of Technology, 2008
Japan Prize, 2013
C. Grant Willson, UT Austin

Motivation for Coarse Graining





30 nm = 600 atoms

$\Rightarrow$ 216,000,000 atoms in a cube

$\Rightarrow$ 216,000,000 x 3 degrees of freedom

= 20 coarse-grained particles

$\Rightarrow$ 8000 particles in a cube

$\Rightarrow$ 24,000 degrees of freedom

Coarse Graining as a Reduced Order Method

- M.L. Huggins, *Journal of Chemical Physics*, 1941
- P.J. Flory, *Journal of Computational Physics*, 1942
- S. Izvekov, M. Parrienllo, C.J. Burnham, and G.A. Voth, *Journal of Chemical Physics*, 2004
- S. Izvekov and G.A. Voth, *Journal of Physical Chemistry B*, 2005, *Journal of Chemical Physics*, 2005, 2006
- W.G. Noid, J.-W. Chu, P. Liu, G.S. Ayton, V. Krishna, S. Izvekov, G.A. Voth, A. Das, and H.C. Anderson, *The Journal of Chemical Physics*, 2008
- J.W. Mullinax and W.G. Noid, *Journal of Chemical Physics*, 2009
- S. Izvekov, P.W. Chung, B.M. Rice, *Journal of Chemical Physics*, 2010
- E. Brini, V. Marcon, and N.F.A. van der Vegt, *Physical Chemistry Chemical Physics*, 2011
- A. Chaimovich and M.S. Shell, *Journal of Chemical Physics*, 2011
- E. Brini and N.F.A. van der Vegt, *Journal of Chemical Physics*, 2012
- S.P. Carmichael and M.S. Shell, *Journal of Physical Chemistry B*, 2012
- Y. Li, B.C. Abberton, M. Kroger, W.K.Liu, *Polymers*, 2013.
- W.G. Noid, *Journal of Chemical Physics*, 2013

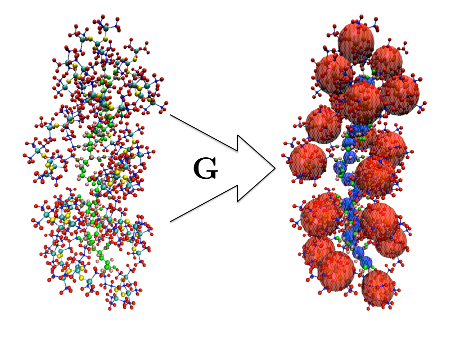
Various CG Methods

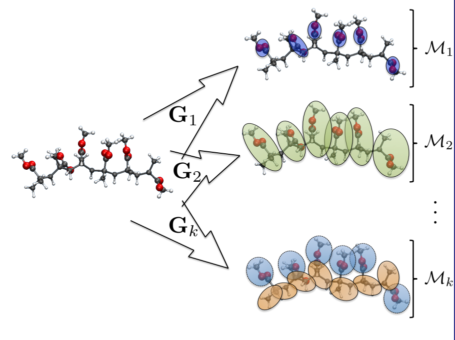
- Force-matching methods
- Multiscale coarse-graining
- Iterative Boltzmann inversion
- Reverse Monte Carlo
- Conditional Reverse Work
- Minimum Relative Entropy

  ⋮

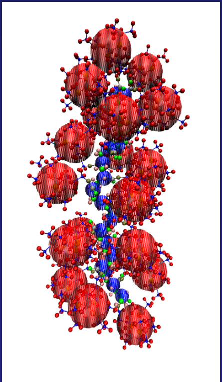
While often advocated, few take into account uncertainties in data, parameters, model inadequacy, . . .

Parametric Model Classes $\mathcal{M}_k$





$$\mathcal{M} = \{\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_k\}$$

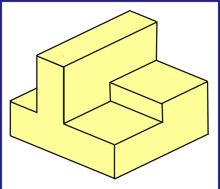


CG Model

$$\frac{\partial U(G(\omega);\boldsymbol{\theta})}{\partial \mathbf{R}_i} - \mathbf{F}_i = \mathbf{0}, \quad i = 1, 2, \ldots, n \qquad (+\text{B.C.}'\text{s})$$

$$U(G(\omega);\boldsymbol{\theta}) = \sum_{i=1}^{N_{co}} \frac{k_i}{2}(\mathbf{R} - \mathbf{R}_{0i})^2 + \sum_{i=1}^{N_\theta} \frac{\kappa_i}{2}(\theta_i - \theta_{0i})^2 + \sum_{i=1}^{N_\omega} \frac{\kappa_i^t}{2}(1 + \cos(i\omega - \gamma))^2 + \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left\{ 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{R_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{R_{ij}}\right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 R_{ij}} \right\}$$

$$\boldsymbol{\theta} = \text{CG model parameters} = \{k_i, \kappa_i, \kappa_i^t, \varepsilon_{ij}, \gamma, \sigma_{ij}\}$$



Macroscale Model

$$\text{Div}\frac{\partial W(\boldsymbol{\mu};\mathbf{w})}{\partial \mathbf{F}} - \mathbf{f} = \mathbf{0}, \quad \mathbf{x} \in \Omega \subset \mathbb{R}^3 \qquad (+\text{B.C.}'\text{s})$$

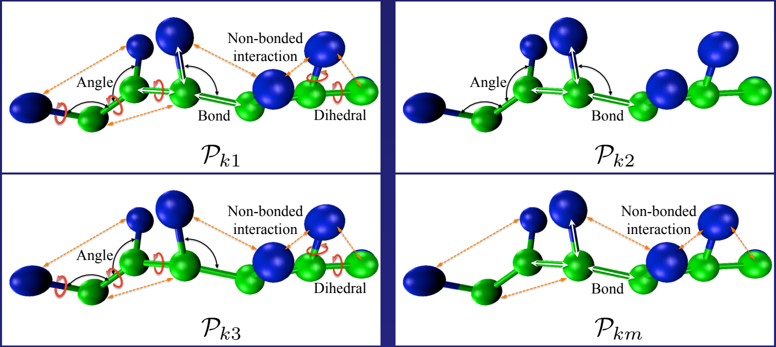
$$W(\boldsymbol{\mu};\mathbf{w}) = \alpha(I_1(\mathbf{C}) - 3) + \beta(I_2(\mathbf{C}) - 3) - \kappa \ln J(\mathbf{C})$$

$$\left(\mathbf{C} = \mathbf{F}^T\mathbf{F}; \quad \mathbf{F} = \mathbf{I} + \boldsymbol{\nabla}\mathbf{w}\right)$$

$$\boldsymbol{\mu} = \text{macromodel parameters} = (\alpha, \beta, \kappa)$$

Parametric Model Classes $\mathcal{M}_k$

$\mathcal{M}_k$:

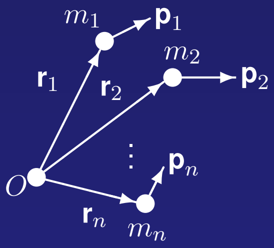


$$\mathcal{M}_i = \{\mathcal{P}_{i1}(\boldsymbol{\theta}_{i1}), \mathcal{P}_{i2}(\boldsymbol{\theta}_{i2}), \dots, \mathcal{P}_{im}(\boldsymbol{\theta}_{im})\}, \qquad i = 1, 2, \dots, k$$

For simplicity in notation:

$$\mathcal{M} = \{\mathcal{P}_1(\boldsymbol{\theta}_1), \mathcal{P}_2(\boldsymbol{\theta}_2), \dots, \mathcal{P}_m(\boldsymbol{\theta}_m)\}$$

What are the Models?

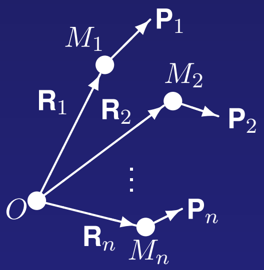


AA Model

$$\mathbf{r}^n = \{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n\}$$

$$\mathbf{R}^N = \{\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_N\}$$

$$\text{“}G(\mathbf{r}^n) = G(\omega) = \mathbf{R}^N\text{”}$$



CG Model

$$G_{A\alpha}\mathbf{r}_\alpha = \mathbf{R}_A; \quad G_{\alpha A}\mathbf{R}_A = \mathbf{r}_\alpha$$

$1 \leq \alpha \leq n$
$1 \leq A \leq N$

Observables

AA $$\langle q \rangle = \int_{\Gamma_{AA}} \rho_{AA}(\mathbf{r}^n) q(\mathbf{r}^n)\, d\mathbf{r}^n = \lim_{\tau \to \infty} \tau^{-1} \int_0^\tau q(\mathbf{r}^n(t))\, dt$$

CG $$Q(\boldsymbol{\theta}) = \int_{\Gamma_{CG}} \rho_{CG}(\mathbf{R}^N, \boldsymbol{\theta}) q(\mathbf{R}^N)\, d\mathbf{R}^N = \lim_{\tau \to \infty} \tau^{-1} \int_0^\tau q(\mathbf{R}^N(t), \boldsymbol{\theta})\, dt$$

What are the Models?

AA

$$m_{\alpha\beta}\ddot{r}_{\beta i} + \frac{\partial}{\partial r_{\alpha i}} u_{AA}(\mathbf{r}^n) - f_{\alpha i} = 0 \qquad \begin{matrix} 1 \le \alpha, \beta \le n \\ 1 \le i \le 3 \end{matrix}$$

CG

$$M_{AB}\ddot{R}_{Bi} + \frac{\partial}{\partial R_{Ai}} U(\mathbf{R}^N, \boldsymbol{\theta}) - F_{Ai} = 0 \qquad \begin{matrix} 1 \le A, B \le N \\ 1 \le i \le 3 \end{matrix}$$

Adjoint

$$m_{\alpha\beta}\ddot{z}_{\beta i} + H_{\alpha i\beta j}(\mathbf{r}^n) z_{\beta j} - \frac{\partial}{\partial r_{\alpha i}} q(\mathbf{r}^n) = 0$$

$$H_{\alpha i\beta j}(\mathbf{r}^n) = \frac{\partial^2 u_{AA}(\mathbf{r}^n)}{\partial r_{\alpha i}\partial r_{\beta j}}$$

What are the Models?

Residual

$$\mathcal{R}(\mathbf{R}^N(\boldsymbol{\theta}),\mathbf{z}^n) = \lim_{\tau\to\infty}\tau^{-1}\int_0^\tau\Big(z_{\alpha i}G_{\alpha A}M_{AB}\ddot{R}_{Bi} + z_{\alpha i}G_{\alpha B}\frac{\partial}{\partial R_{Bi}}U(\mathbf{R}^N,\boldsymbol{\theta}) - z_{\alpha i}G_{\alpha B}F_{Bi}\Big)\,dt$$

Theorem

(Under suitable smoothness conditions), the error in the observables due to the CG approximation is, $\forall\boldsymbol{\theta}\in\Theta$,

$$\varepsilon(\boldsymbol{\theta}) = \langle q\rangle - Q(\boldsymbol{\theta}) = \mathcal{R}(\mathbf{R}^N(\boldsymbol{\theta}),\mathbf{z}^n) + \Delta \approx \mathcal{R}(\mathbf{R}^N(\boldsymbol{\theta}),\mathbf{z}^n)$$

where Δ is a remainder of higher order in "$\left\|\mathbf{r}^n - \mathbf{R}^N\right\|$"

Information Entropy

Suppose

$$\begin{aligned} Q(\mathbf{r}^n) &= \int_\Gamma \rho(\mathbf{r}^n)\, q(\mathbf{r}^n)\; d\mathbf{r}^n \\ Q\left(\mathbf{R}^N(\boldsymbol{\theta})\right) &= \int_\Gamma \rho(\mathbf{r}^n)\, q\left(G\left(\mathbf{r}^n(\boldsymbol{\theta})\right)\right)\; d\mathbf{r}^n \\ q(\mathbf{r}^n) &= \log \rho(\mathbf{r}^n) \end{aligned}$$

Then

$$\begin{aligned} Q(\mathbf{r}^n) - Q\left(\mathbf{R}^N(\boldsymbol{\theta})\right) &= D_{KL}\left(\rho(\mathbf{r}^n) \,\|\, \rho\left(G\left(\mathbf{r}^n(\boldsymbol{\theta})\right)\right)\right) \\ &= \mathcal{R}\left(\mathbf{R}^N(\boldsymbol{\theta}), \mathbf{z}^n\right) \end{aligned}$$

Information Entropy

Suppose

$$
\begin{aligned}
Q\left(\mathbf{r}^n\right) &= \int_\Gamma \rho\left(\mathbf{r}^n\right) q\left(\mathbf{r}^n\right)\, d\mathbf{r}^n \\
Q\left(\mathbf{R}^N\left(\boldsymbol{\theta}\right)\right) &= \int_\Gamma \rho\left(\mathbf{r}^n\right) q\left(G\left(\mathbf{r}^n\left(\boldsymbol{\theta}\right)\right)\right)\, d\mathbf{r}^n \\
q\left(\mathbf{r}^n\right) &= \log \rho\left(\mathbf{r}^n\right)
\end{aligned}
$$

Then

$$
\begin{aligned}
Q\left(\mathbf{r}^n\right) - Q\left(\mathbf{R}^N\left(\boldsymbol{\theta}\right)\right) &= D_{KL}\left(\rho\left(\mathbf{r}^n\right) \| \rho\left(G\left(\mathbf{r}^n\left(\boldsymbol{\theta}\right)\right)\right)\right) \\
&= \mathcal{R}\left(\mathbf{R}^N\left(\boldsymbol{\theta}\right), \mathbf{z}^n\right)
\end{aligned}
$$

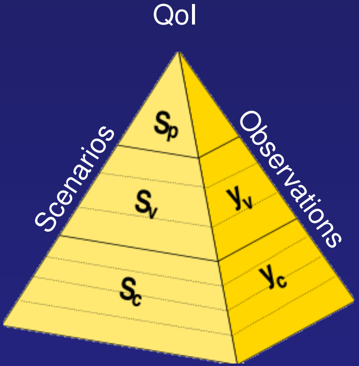
$$
= \int \rho(\omega) \log \frac{\rho(\omega)}{\rho(G(\omega))}\, d\omega
$$

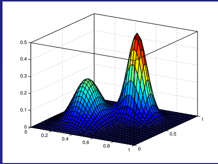
3. Bayesian Model Calibration, Validation, and Prediction

"The essence of 'honesty' or 'objectivity' demands that we take into account all of the evidence we have, not just some arbitrarily chosen subset of it."

—E.T. Jaynes, 2003

Climbing the Prediction Pyramid





Prior

$$\pi(\boldsymbol{\theta})$$

Calibration $(S_c, \mathbf{y}_c)$

$$\pi(\boldsymbol{\theta}|\mathbf{y}_c) = \frac{\pi(\mathbf{y}_c|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{y}_c)}$$

Validation $(S_v, \mathbf{y}_v)$

$$\pi(\boldsymbol{\theta}|\mathbf{y}_v, \mathbf{y}_c) = \frac{\pi(\mathbf{y}_v|\boldsymbol{\theta}, \mathbf{y}_c)\pi(\boldsymbol{\theta}, \mathbf{y}_c)}{\pi(\mathbf{y}_v|\mathbf{y}_c)}$$

Prediction $(S_p$, QoI)

$$\pi(Q) = \pi(Q|\boldsymbol{\theta}, S_v, S_c)$$

Basic Ideas:

- Use statistical inverse methods based on Bayes' rule to calibrate parameters

$$\pi(\boldsymbol{\theta}|\mathbf{y}_c) = \frac{\pi(\mathbf{y}_c|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{y}_c)}$$

- Design validation experiments to challenge model assumptions and inform model of QoIs

$$\pi(\boldsymbol{\theta}|\mathbf{y}_v, \mathbf{y}_c) = \frac{\pi(\mathbf{y}_v|\boldsymbol{\theta}, \mathbf{y}_c)\pi(\boldsymbol{\theta}|\mathbf{y}_c)}{\pi(\mathbf{y}_v|\mathbf{y}_c)}$$

- Is model "valid" (not invalid) for the validation QoI (observable) given the data and predictions $\pi(Q_{vk}|\mathbf{y}_{vk})$, $\pi(Q|\mathbf{y}_c)$?
- Solve forward problem for the QoI (not observable) using validation parameters

$$\pi(Q) = \int \pi(Q|\boldsymbol{\theta}, \mathbf{y}_{vk}, \mathbf{y}_{vk-1}, \ldots, \mathbf{y}_c)\, d\boldsymbol{\theta}$$

- Compute quantity of uncertainty in $\pi(Q)$

Basic Ideas:

- Use statistical inverse methods based on Bayes' rule to calibrate parameters

$$\pi(\boldsymbol{\theta}|\mathbf{y}_c) = \frac{\pi(\mathbf{y}_c|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{y}_c)}$$

What is the likelihood function? What are the priors? How does one compute the posterior?

- Design validation experiments to challenge model assumptions and inform model of QoIs

$$\pi(\boldsymbol{\theta}|\mathbf{y}_v, \mathbf{y}_c) = \frac{\pi(\mathbf{y}_v|\boldsymbol{\theta}, \mathbf{y}_c)\pi(\boldsymbol{\theta}|\mathbf{y}_c)}{\pi(\mathbf{y}_v|\mathbf{y}_c)}$$

- Is model "valid" (not invalid) for the validation QoI (observable) given the data and predictions $\pi(Q_{vk}|\mathbf{y}_{vk})$, $\pi(Q|\mathbf{y}_c)$?
- Solve forward problem for the QoI (not observable) using validation parameters

$$\pi(Q) = \int \pi(Q|\boldsymbol{\theta}, \mathbf{y}_{vk}, \mathbf{y}_{vk-1}, \ldots, \mathbf{y}_c)\, d\boldsymbol{\theta}$$

- Compute quantity of uncertainty in $\pi(Q)$

Basic Ideas:

- Use statistical inverse methods based on Bayes' rule to calibrate parameters

$$\pi(\boldsymbol{\theta}|\mathbf{y}_c) = \frac{\pi(\mathbf{y}_c|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{y}_c)}$$

What is the likelihood function? What are the priors? How does one compute the posterior?

- Design validation experiments to challenge model assumptions and inform model of QoIs

$$\pi(\boldsymbol{\theta}|\mathbf{y}_v, \mathbf{y}_c) = \frac{\pi(\mathbf{y}_v|\boldsymbol{\theta}, \mathbf{y}_c)\pi(\boldsymbol{\theta}|\mathbf{y}_c)}{\pi(\mathbf{y}_v|\mathbf{y}_c)}$$

Is the validation experiment well chosen? Has it resulted in an information gain over the calibration?

- Is model "valid" (not invalid) for the validation QoI (observable) given the data and predictions $\pi(Q_{vk}|\mathbf{y}_{vk})$, $\pi(Q|\mathbf{y}_c)$?
- Solve forward problem for the QoI (not observable) using validation parameters

$$\pi(Q) = \int \pi(Q|\boldsymbol{\theta}, \mathbf{y}_{vk}, \mathbf{y}_{vk-1}, \ldots, \mathbf{y}_c)\, d\boldsymbol{\theta}$$

- Compute quantity of uncertainty in $\pi(Q)$

Basic Ideas:

- Use statistical inverse methods based on Bayes' rule to calibrate parameters

$$\pi(\boldsymbol{\theta}|\mathbf{y}_c) = \frac{\pi(\mathbf{y}_c|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{y}_c)}$$

What is the likelihood function? What are the priors? How does one compute the posterior?

- Design validation experiments to challenge model assumptions and inform model of QoIs

$$\pi(\boldsymbol{\theta}|\mathbf{y}_v, \mathbf{y}_c) = \frac{\pi(\mathbf{y}_v|\boldsymbol{\theta}, \mathbf{y}_c)\pi(\boldsymbol{\theta}|\mathbf{y}_c)}{\pi(\mathbf{y}_v|\mathbf{y}_c)}$$

Is the validation experiment well chosen? Has it resulted in an information gain over the calibration?

- Is model "valid" (not invalid) for the validation QoI (observable) given the data and predictions $\pi(Q_{vk}|\mathbf{y}_{vk})$, $\pi(Q|\mathbf{y}_c)$?

What is the criterion for "validity" of a model?

- Solve forward problem for the QoI (not observable) using validation parameters

$$\pi(Q) = \int \pi(Q|\boldsymbol{\theta}, \mathbf{y}_{vk}, \mathbf{y}_{vk-1}, \ldots, \mathbf{y}_c)\, d\boldsymbol{\theta}$$

- Compute quantity of uncertainty in $\pi(Q)$

Basic Ideas:

- Use statistical inverse methods based on Bayes' rule to calibrate parameters

$$\pi(\boldsymbol{\theta}|\mathbf{y}_c) = \frac{\pi(\mathbf{y}_c|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{y}_c)}$$

What is the likelihood function? What are the priors? How does one compute the posterior?

- Design validation experiments to challenge model assumptions and inform model of QoIs

$$\pi(\boldsymbol{\theta}|\mathbf{y}_v, \mathbf{y}_c) = \frac{\pi(\mathbf{y}_v|\boldsymbol{\theta}, \mathbf{y}_c)\pi(\boldsymbol{\theta}|\mathbf{y}_c)}{\pi(\mathbf{y}_v|\mathbf{y}_c)}$$

Is the validation experiment well chosen? Has it resulted in an information gain over the calibration?

- Is model "valid" (not invalid) for the validation QoI (observable) given the data and predictions $\pi(Q_{vk}|\mathbf{y}_{vk})$, $\pi(Q|\mathbf{y}_c)$?

What is the criterion for "validity" of a model?

- Solve forward problem for the QoI (not observable) using validation parameters

$$\pi(Q) = \int \pi(Q|\boldsymbol{\theta}, \mathbf{y}_{vk}, \mathbf{y}_{vk-1}, \ldots, \mathbf{y}_c)\, d\boldsymbol{\theta}$$

How do we solve the forward problem?

- Compute quantity of uncertainty in $\pi(Q)$

Basic Ideas:

- Use statistical inverse methods based on Bayes' rule to calibrate parameters

$$\pi(\boldsymbol{\theta}|\mathbf{y}_c) = \frac{\pi(\mathbf{y}_c|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{y}_c)}$$

What is the likelihood function? What are the priors? How does one compute the posterior?

- Design validation experiments to challenge model assumptions and inform model of QoIs

$$\pi(\boldsymbol{\theta}|\mathbf{y}_v, \mathbf{y}_c) = \frac{\pi(\mathbf{y}_v|\boldsymbol{\theta}, \mathbf{y}_c)\pi(\boldsymbol{\theta}|\mathbf{y}_c)}{\pi(\mathbf{y}_v|\mathbf{y}_c)}$$

Is the validation experiment well chosen? Has it resulted in an information gain over the calibration?

- Is model "valid" (not invalid) for the validation QoI (observable) given the data and predictions $\pi(Q_{vk}|\mathbf{y}_{vk})$, $\pi(Q|\mathbf{y}_c)$?

What is the criterion for "validity" of a model?

- Solve forward problem for the QoI (not observable) using validation parameters

$$\pi(Q) = \int \pi(Q|\boldsymbol{\theta}, \mathbf{y}_{vk}, \mathbf{y}_{vk-1}, \ldots, \mathbf{y}_c)\, d\boldsymbol{\theta}$$

How do we solve the forward problem?

- Compute quantity of uncertainty in $\pi(Q)$

How do we "quantify" uncertainty?

The Prior

We seek a logical measure $H(p)$ of the amount of uncertainty in a probability distribution $p = \{p_1, p_2, \ldots, p_n\}$, $p_i = p(x_i)$

1. $H(p) \in \mathbb{R}$
2. $H \in C^0(\mathbb{R})$
3. "common sense:" $H(\frac{1}{n}, \frac{1}{n}, \ldots) \uparrow$ as $n \to \infty$
4. Consistency

Shannon's Theorem

The only function satisfying four logical desiderata is the information entropy

$$H(p) = -\sum_{i=1}^{n} p_i \log p_i \qquad \left(\text{or} \quad -\int p \log p/m \, dx\right)$$

Moreover, the actual probability p *maximizes* $H(p)$ subject to constraints imposed by available information

Relative Entropy

Given two pdfs p and q, the relative entropy is given by the Kullback-Leibler divergence,

$$D_{KL}(p \| q) = \int p(x) \log \frac{p(x)}{q(x)} \, dx = -H(p) + H(p, q)$$

The Prior

Maximize $H(p)$ subject to prior information constraints:

- $\langle x \rangle$

$$\mathcal{L}(p,\lambda) = H(p) - \lambda_0 \left(\sum_{i=1}^{n} p_i - 1 \right) - \lambda_1 \left(\sum_{i=1}^{n} p_i x_i - \langle x \rangle \right)$$

$$\Rightarrow \qquad \boxed{\pi(\boldsymbol{\theta}) = \frac{1}{\langle x \rangle} \exp\left\{ -x / \langle x \rangle \right\}}$$

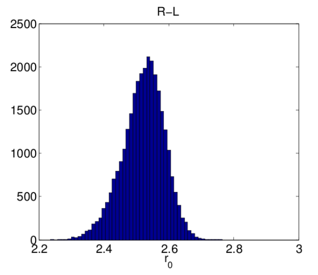
- $\langle x \rangle, \sigma_x^2$

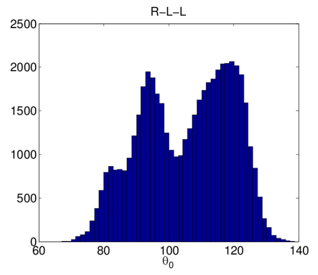
$$\begin{aligned} \mathcal{L}(p,\lambda) &= H(p) - \lambda_0 \left(\sum_{i=1}^{n} p_i - 1 \right) - \lambda_1 \left(\sum_{i=1}^{n} p_i x_i - \langle x \rangle \right) \\ &\quad - \lambda_2 \left(\sum_{i=1}^{n} p_i \left(x_i - \langle x \rangle \right)^2 - \sigma_x^2 \right) \end{aligned}$$

$$\Rightarrow \qquad \boxed{\pi(\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left\{ \frac{-(x - \langle x \rangle)^2}{2\sigma_x^2} \right\}}$$

E. T. Jaynes (1988)

Determining Calibration Priors: Bonds





Bond Equilibrium Distance: R_0

$\langle R_0 \rangle = 2.5219$

$\sigma^2_{R_0} = 4.1097 \times 10^{-3}$

Spring Constant: k_{R_0}

$\langle k_R \rangle = k_B T / 2\sigma^2_{R_0}$

$= 72.5264$

Equilibrium Angle: $\theta_0,$

$\langle \theta_0 \rangle = 105.5117$

$\sigma^2_{\theta_0} = 192.8262$

Spring Constant: k_θ

$\langle k_\theta \rangle = k_B T / 2\sigma^2_{\theta_0}$

$= 1.5458 \times 10^{-3}$

The Likelihood Function

R.A. Fisher, 1922: The likelihood that any parameter should have any assigned value is proportional to the probability that if this were true the totality of all observations should be that observed.

Consider n i.i.d. random observables $y_1, y_2, \ldots, y_n$

For each sample,

$$\pi(y_i|\boldsymbol{\theta}) = p(y_i - d_i(\boldsymbol{\theta}))$$

For many samples,

$$\pi(y_1, y_2, \ldots, y_n|\boldsymbol{\theta}) = \prod_{i=1}^{n} \pi(y_i|\boldsymbol{\theta})$$

Then the log-likelihood is

$$L_n(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log \pi(y_i|\boldsymbol{\theta})$$

The Model Evidence - Model Plausibilities: Which model is "best"?

$\mathcal{M}$ = set of parametric model classes $= \{\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_m\}$

Each $\mathcal{P}$ has its own likelihood and parameters $\boldsymbol{\theta}_j$

Bayes' rule in expanded form:

$$\pi(\boldsymbol{\theta}_j|\mathbf{y},\mathcal{P}_j,\mathcal{M}) = \frac{\pi(\mathbf{y}|\boldsymbol{\theta}_j,\mathcal{P}_j,\mathcal{M})\pi(\boldsymbol{\theta}_j|\mathcal{P}_j,\mathcal{M})}{\pi(\mathbf{y}|\mathcal{P}_j,\mathcal{M})}, \quad 1 \le j \le m$$

$$\text{model evidence} = \int \pi(\mathbf{y}|\boldsymbol{\theta}_j,\mathcal{P}_j,\mathcal{M})\pi(\boldsymbol{\theta}_j|\mathcal{P}_j,\mathcal{M})\, d\boldsymbol{\theta}_j$$

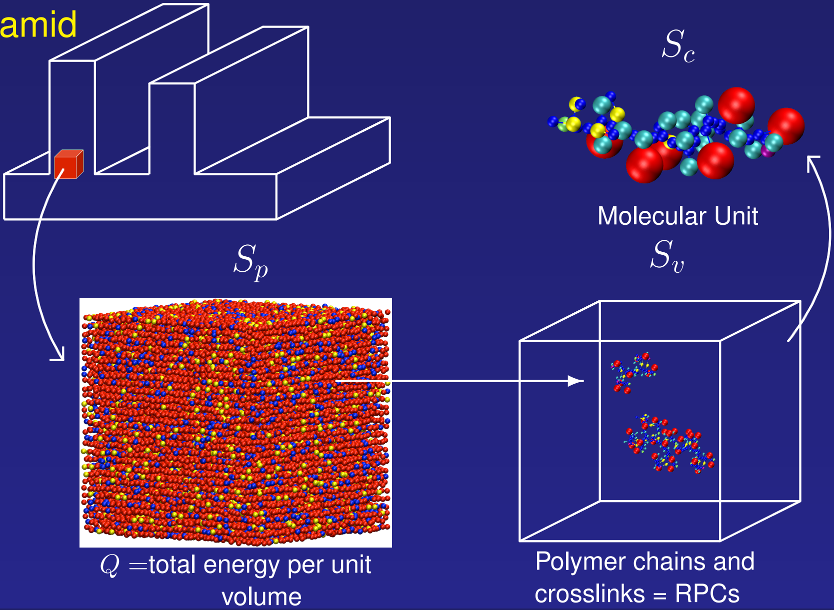
Now apply Bayes' rule to the evidence:

$$\rho_j = \pi(\mathcal{P}_j|\mathbf{y},\mathcal{M}) = \frac{\pi(\mathcal{P}_j|\mathcal{M})}{\pi(\mathbf{y}|\mathcal{M})}\pi(\mathbf{y}|\mathcal{P}_j,\mathcal{M})$$

$=$ the *posterior model plausibility*

$$\sum_{j=1}^{m}\rho_j = \frac{1}{\pi(\mathbf{y}|\mathcal{M})}\sum_{j=1}^{m}\pi(\mathbf{y}|\mathcal{P}_j,\mathcal{M})\pi(\mathcal{P}_j|\mathcal{M}) = 1$$

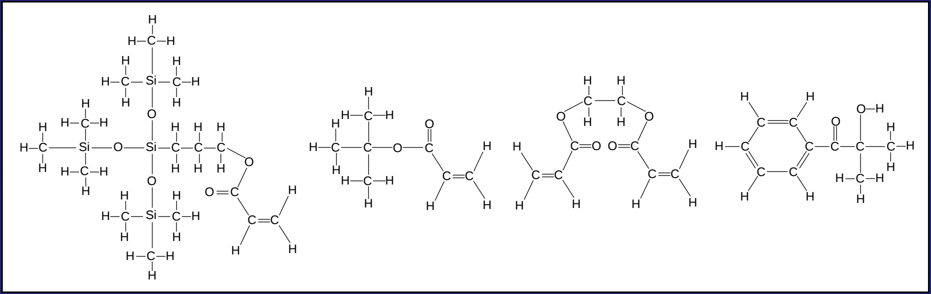
4. The Prediction Process: Traveling up the Prediction Pyramid



SFIL Coarse Graining

Constituents of Etch Barrier

Monomer 1 Monomer 2 Crosslinker Initiator

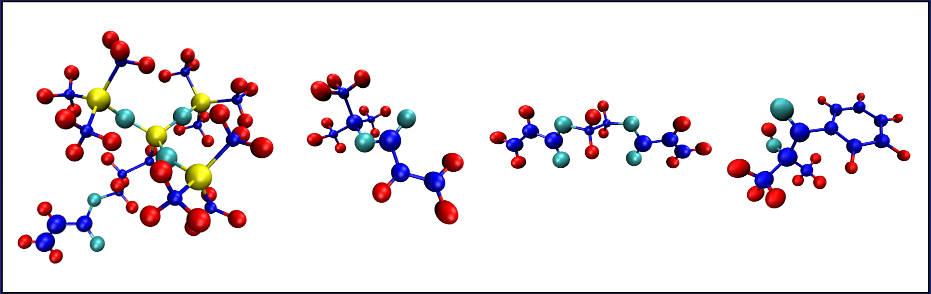


Farrell and Oden 2013

SFIL Coarse Graining

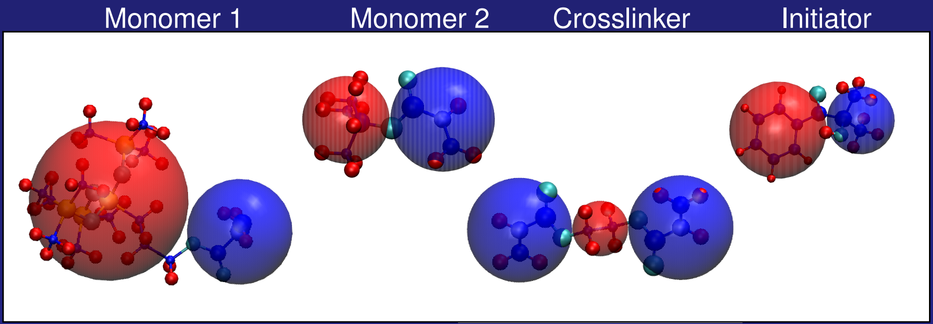
Constituents of Etch Barrier

Monomer 1 Monomer 2 Crosslinker Initiator

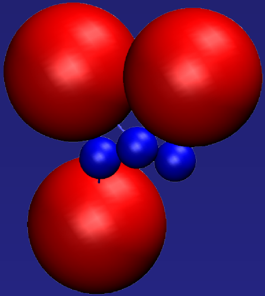


SFIL Coarse Graining

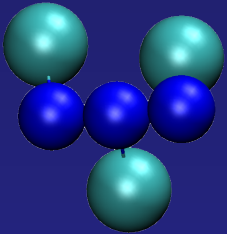
Constituents of Etch Barrier



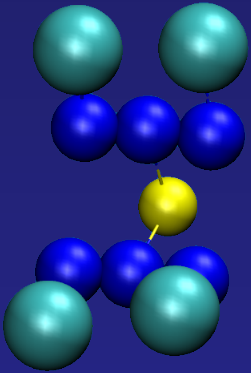
SFIL Calibration Scenarios: S_c



S_{c1}

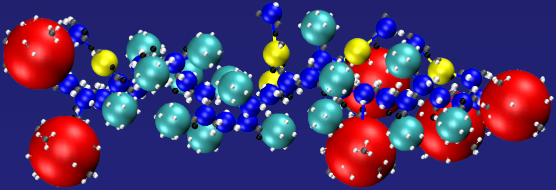


S_{c2}

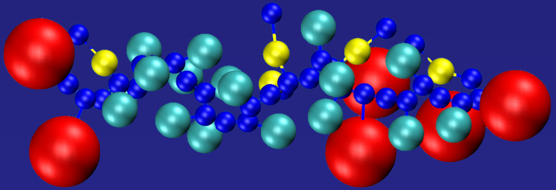


S_{c3}

SFIL Coarse Graining



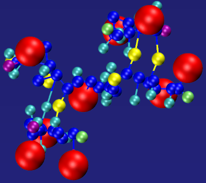
AA System
827 atoms
503 parameters



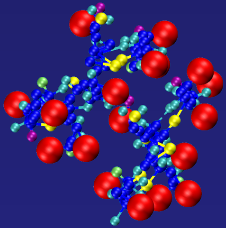
CG System
61 particles

SFIL Validation Scenario: S_v

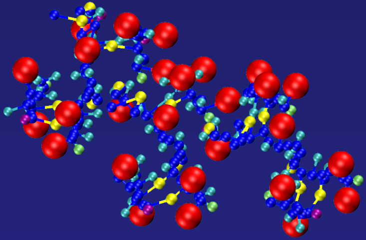
Series of scenarios increasing in size



$S_{v,1}$



$S_{v,2}$



$S_{v,3}$

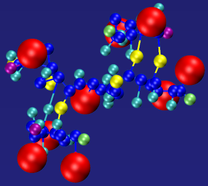
For each scenario compute the QoI:

$$Q = \int_{\Gamma_{AA}} \rho\left(\mathbf{r}^n\right) u_{AA}\left(\mathbf{r}^n\right)\, d\mathbf{r}^n; \quad Q_{v,k}\left(\boldsymbol{\theta}\right) = \int_{\Gamma_{CG,k}} \rho\left(\mathbf{R}^N\right) U_{CG}\left(\mathbf{R}^N; \boldsymbol{\theta}\right)\, d\mathbf{R}^N$$

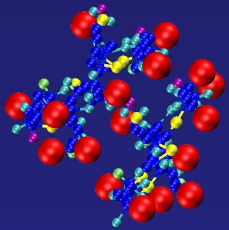
$$\rho\left(\mathbf{r}^n\right) \propto \exp\left\{-\beta u\left(\mathbf{r}^n\right)\right\}$$

SFIL Validation Scenario: S_v

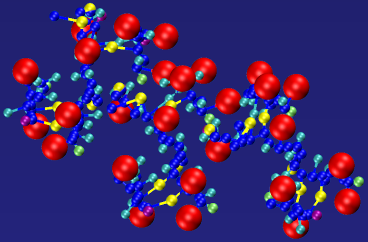
Series of scenarios increasing in size



$S_{v,1}$



$S_{v,2}$



$S_{v,3}$

Compare the computed QoI to AA data: if

$$D_{KL}\left(\pi(u_{AA}|_{S_v})\|\pi(Q_v)\right) < \gamma_{tol}$$

the model is considered validated

SFIL Model Classes

Model	Bonds	Angles	Dihedrals	Non-Bonded	# of Parameters
$\mathcal{P}_1$				✓	12
$\mathcal{P}_2$	✓				18
$\mathcal{P}_3$	✓			✓	30
$\mathcal{P}_4$		✓			32
$\mathcal{P}_5$		✓		✓	44
$\mathcal{P}_6$	✓	✓			50
$\mathcal{P}_7$	✓	✓		✓	62
$\mathcal{P}_8$			✓		96
$\mathcal{P}_9$			✓	✓	108
$\mathcal{P}_{10}$	✓		✓		114
$\mathcal{P}_{11}$	✓		✓	✓	126
$\mathcal{P}_{12}$		✓	✓		128
$\mathcal{P}_{13}$		✓	✓	✓	140
$\mathcal{P}_{14}$	✓	✓	✓		146
$\mathcal{P}_{15}$	✓	✓	✓	✓	158

Sensitivity Analysis

- PIRT (Phenomena Identification and Ranking Table)
- Importance Measures (Hora and Iman, 1995)
- Correlation Ratios (McKay, 1995)
- Sensitivity Analysis (Saltelli, Chan, Scott, 2000, Saltelli *et. al.* 2008)
- Variance-based

$$S_{i_1,\dots,i_k} = \frac{V_{i_1,\dots,i_k}(Y)}{V(Y)}$$

- Entropy-based

$$\begin{aligned} KL_i(p_1 \| p_0) &= \int_{-\infty}^{\infty} p_1(y(\theta_1, \theta_2, \dots, \bar{\theta}_i, \dots, \theta_m)) \\ &\quad \times \left| \log \frac{p_0(y(\theta_1, \theta_2, \dots, \bar{\theta}_i, \dots, \theta_m))}{p_1(y(\theta_1, \theta_2, \dots, \theta_i, \dots, \theta_m))} \right| \, dy, \quad \bar{\theta}_i = \langle \theta_i \rangle \\ &= D_{KL} \end{aligned}$$

- Scatter Plots, etc

Saltelli, A., $et.al.$ (2001)
Auder, B. and Iooss, B. (2009)
Chen, W $et.al$ (2005)

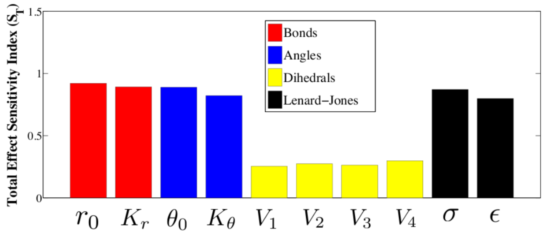
Sensitivity Analysis: Variance-Based Method

$Y(\boldsymbol{\theta})$ = model output (*e.g.* $Y(\boldsymbol{\theta}) = \langle U(\cdot;\boldsymbol{\theta})\rangle_{CG}$)
$V(Y)$ = output variance $= \mathbb{E}(Y^2) - \mathbb{E}^2(Y)$

$$V(Y) = V_{\boldsymbol{\theta}_{\sim i}}[(\mathbb{E}_{X_i}(Y|\boldsymbol{\theta}_{\sim i}))] + \mathbb{E}_{\boldsymbol{\theta}_{\sim i}}[(V_{X_i}(Y|\boldsymbol{\theta}_{\sim i}))]$$

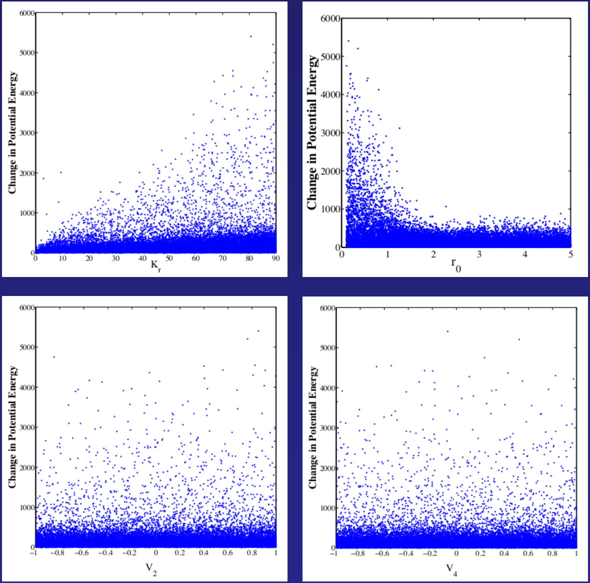
$$S_{T_i} = 1 - \frac{V_{\boldsymbol{\theta}_{\sim i}}\left[\mathbb{E}_{\theta_i}(Y|\boldsymbol{\theta}_{\sim i})\right]}{V(Y)}$$ (Total effect sensitivity index)

Example: Polyethylene chain with 24 beads



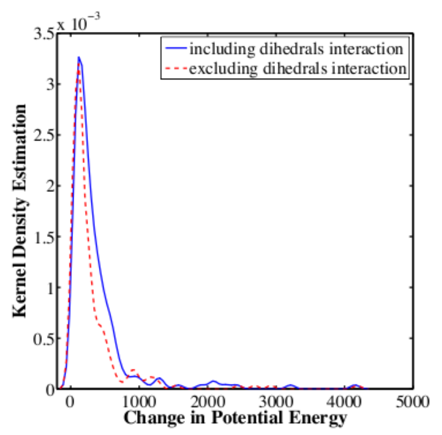
Saltelli, A., *et.al.* (2001)

Sensitivity Analysis: Monte Carlo Scatterplots



Sensitivity Analysis: Comparison

The sensitivity indices and scatterplots show that the dihedral parameters are unimportant, but how important are they?



Occam's Razor

Principle of Occam's Razor

Among competing theories that lead to the same prediction, the one that relies on the fewest assumptions is the best.

When choosing among a set of models:

The simplest valid model is the best choice.

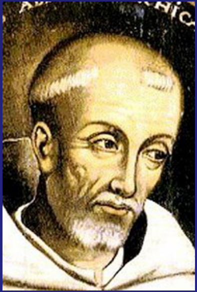
Occam's Razor

Principle of Occam's Razor

Among competing theories that lead to the same prediction, the one that relies on the fewest assumptions is the best.

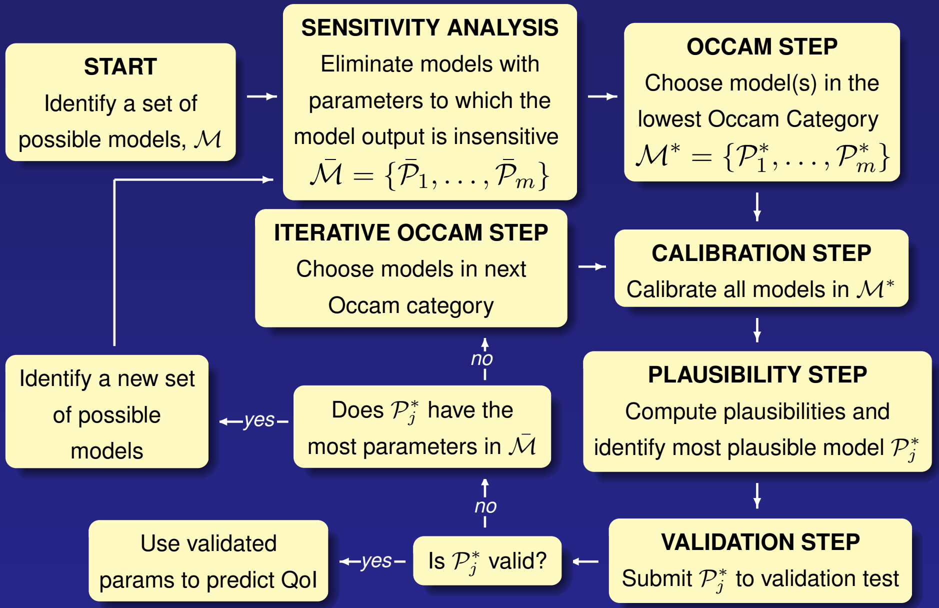
When choosing among a set of models:
The simplest valid model is the best choice.

- simple ⇒ number of parameters
- valid ⇒ passes Bayesian validation test



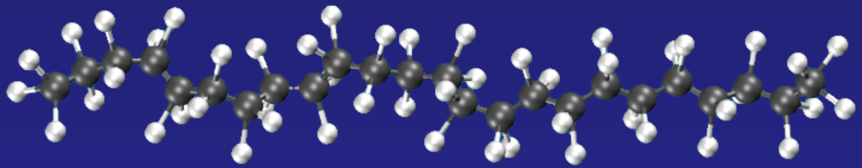
How do we choose a model that adheres to this principle?

The Occam-Plausibility Algorithm



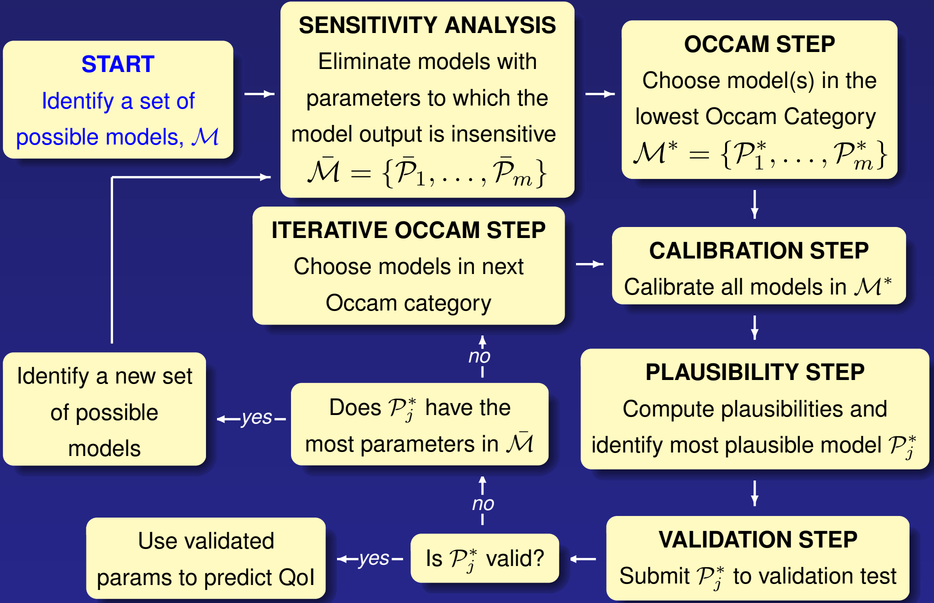
5. Exploratory Example: Polyethylene

Consider as an example polyethylene



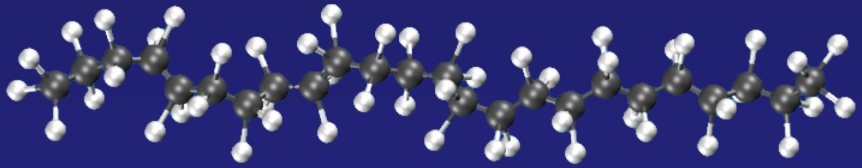
$C_{24}H_{50}$

The Occam-Plausibility Algorithm

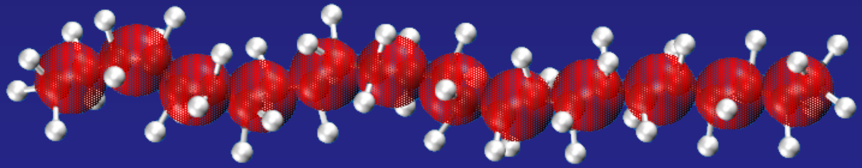


Example: Occam-Plausibility Algorithm

Consider as an example polyethylene



Define the coarse-grained map: 2 carbons per bead

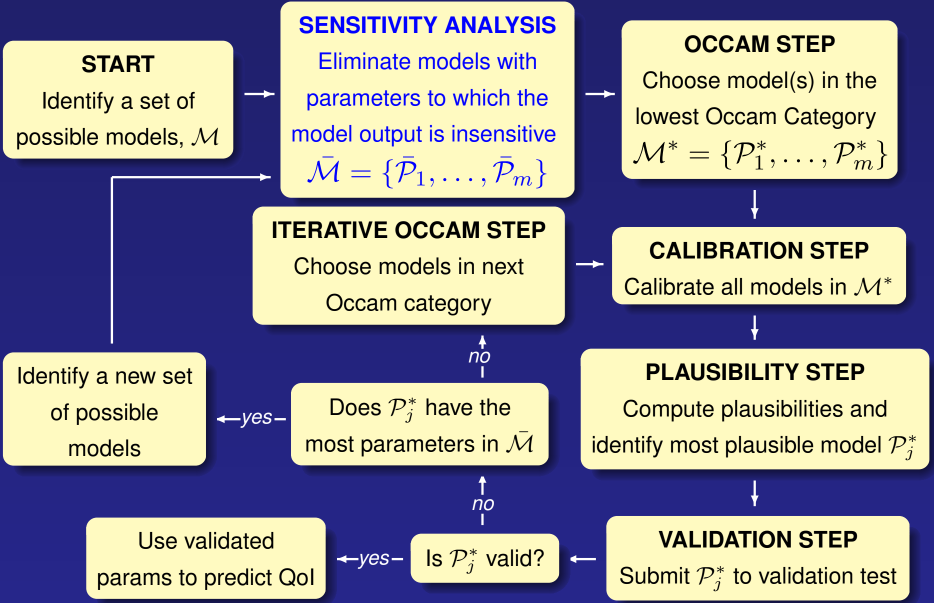


Example: Occam-Plausibility Algorithm (cont)

Representation of the CG potential using OPLS form

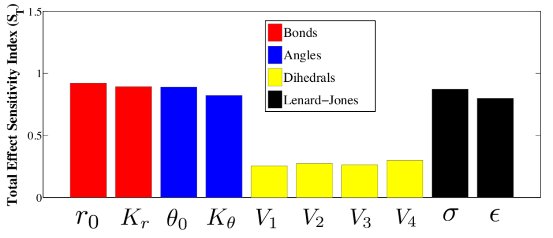
Model	Bonds	Angles	Dihedrals	Non-Bonded	Params
$\mathcal{P}_1$	✓				2
$\mathcal{P}_2$		✓			2
$\mathcal{P}_3$				✓	2
$\mathcal{P}_4$	✓	✓			4
$\mathcal{P}_5$	✓			✓	4
$\mathcal{P}_6$		✓		✓	4
$\mathcal{P}_7$	✓	✓		✓	6
$\mathcal{P}_8$			✓		4
$\mathcal{P}_9$	✓		✓		6
$\mathcal{P}_{10}$		✓	✓		6
$\mathcal{P}_{11}$			✓	✓	6
$\mathcal{P}_{12}$	✓	✓	✓		8
$\mathcal{P}_{13}$	✓		✓	✓	8
$\mathcal{P}_{14}$		✓	✓	✓	8
$\mathcal{P}_{15}$	✓	✓	✓	✓	10

The Occam-Plausibility Algorithm



Example: Occam-Plausibility Algorithm (cont)

$Y = \langle U(\cdot;\boldsymbol{\theta})\rangle$ = potential energy



Example: Occam-Plausibility Algorithm (cont)

Model	Bonds	Angles	Dihedrals	Non-Bonded	Params
$\mathcal{P}_1$	✓				2
$\mathcal{P}_2$		✓			2
$\mathcal{P}_3$				✓	2
$\mathcal{P}_4$	✓	✓			4
$\mathcal{P}_5$	✓			✓	4
$\mathcal{P}_6$		✓		✓	4
$\mathcal{P}_7$	✓	✓		✓	6
$\mathcal{P}_8$			✓		4
$\mathcal{P}_9$	✓		✓		6
$\mathcal{P}_{10}$		✓	✓		6
$\mathcal{P}_{11}$			✓	✓	6
$\mathcal{P}_{12}$	✓	✓	✓		8
$\mathcal{P}_{13}$	✓		✓	✓	8
$\mathcal{P}_{14}$		✓	✓	✓	8
$\mathcal{P}_{15}$	✓	✓	✓	✓	10

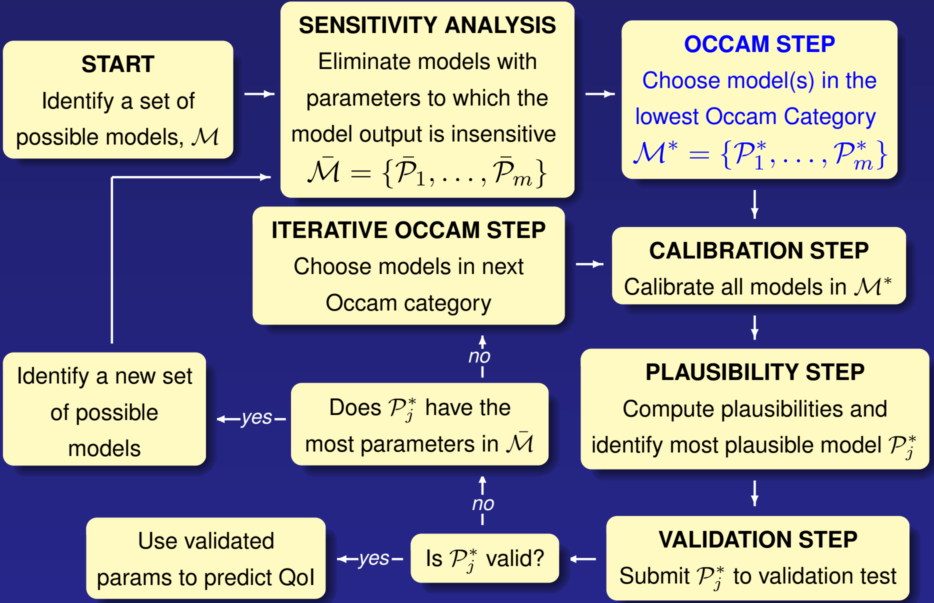
Example: Occam-Plausibility Algorithm (cont)

Model	Bonds	Angles	Dihedrals	LJ 12-6	LJ 9-6	Params
$\bar{\mathcal{P}}_1$	✓					2
$\bar{\mathcal{P}}_2$		✓				2
$\bar{\mathcal{P}}_3$				✓		2
$\bar{\mathcal{P}}_4$					✓	2
$\bar{\mathcal{P}}_5$	✓	✓				4
$\bar{\mathcal{P}}_6$	✓			✓		4
$\bar{\mathcal{P}}_7$	✓				✓	4
$\bar{\mathcal{P}}_8$		✓		✓		4
$\bar{\mathcal{P}}_9$		✓			✓	4
$\bar{\mathcal{P}}_{10}$	✓	✓		✓		6
$\bar{\mathcal{P}}_{11}$	✓			✓	✓	6

Example: Occam-Plausibility Algorithm (cont)

Model	Bonds	Angles	Dihedrals	LJ 12-6	LJ 9-6	Params	Category
$\bar{\mathcal{P}}_1$	✓					2	1
$\bar{\mathcal{P}}_2$		✓				2	
$\bar{\mathcal{P}}_3$				✓		2	
$\bar{\mathcal{P}}_4$					✓	2	
$\bar{\mathcal{P}}_5$	✓	✓				4	2
$\bar{\mathcal{P}}_6$	✓			✓		4	
$\bar{\mathcal{P}}_7$	✓				✓	4	
$\bar{\mathcal{P}}_8$		✓		✓		4	
$\bar{\mathcal{P}}_9$		✓			✓	4	
$\bar{\mathcal{P}}_{10}$	✓	✓		✓		6	3
$\bar{\mathcal{P}}_{11}$	✓			✓	✓	6	

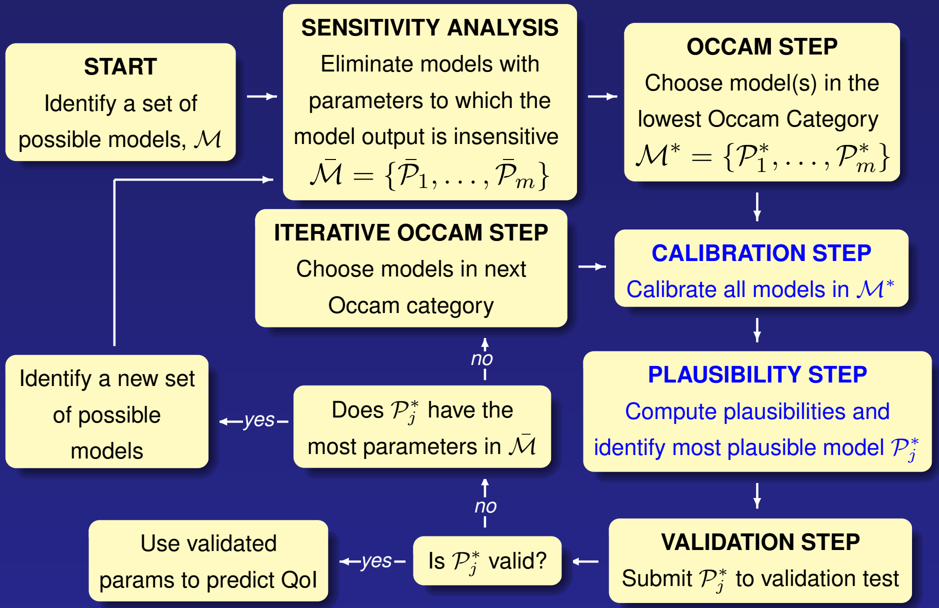
The Occam-Plausibility Algorithm



Example: Occam-Plausibility Algorithm (cont)

Model	Bonds	Angles	Dihedrals	LJ 12-6	LJ 9-6	Params	Category
$\bar{\mathcal{P}}_1$	✓					2	1
$\bar{\mathcal{P}}_2$		✓				2	
$\bar{\mathcal{P}}_3$				✓		2	
$\bar{\mathcal{P}}_4$					✓	2	
$\bar{\mathcal{P}}_5$	✓	✓				4	2
$\bar{\mathcal{P}}_6$	✓			✓		4	
$\bar{\mathcal{P}}_7$	✓				✓	4	
$\bar{\mathcal{P}}_8$		✓		✓		4	
$\bar{\mathcal{P}}_9$		✓			✓	4	
$\bar{\mathcal{P}}_{10}$	✓	✓		✓		6	3
$\bar{\mathcal{P}}_{11}$	✓			✓	✓	6	

The Occam-Plausibility Algorithm



Example: Occam-Plausibility Algorithm (cont)

Calibration

$$\pi(\boldsymbol{\theta}_j^*|\mathbf{y},\mathcal{P}_j^*,\mathcal{M}^*) = \frac{\pi(\mathbf{y}|\boldsymbol{\theta}_j^*,\mathcal{P}_j^*,\mathcal{M}^*)\pi(\boldsymbol{\theta}_j^*|\mathcal{P}_j^*,\mathcal{M}^*)}{\pi(\mathbf{y}|\mathcal{P}_j^*,\mathcal{M}^*)}$$

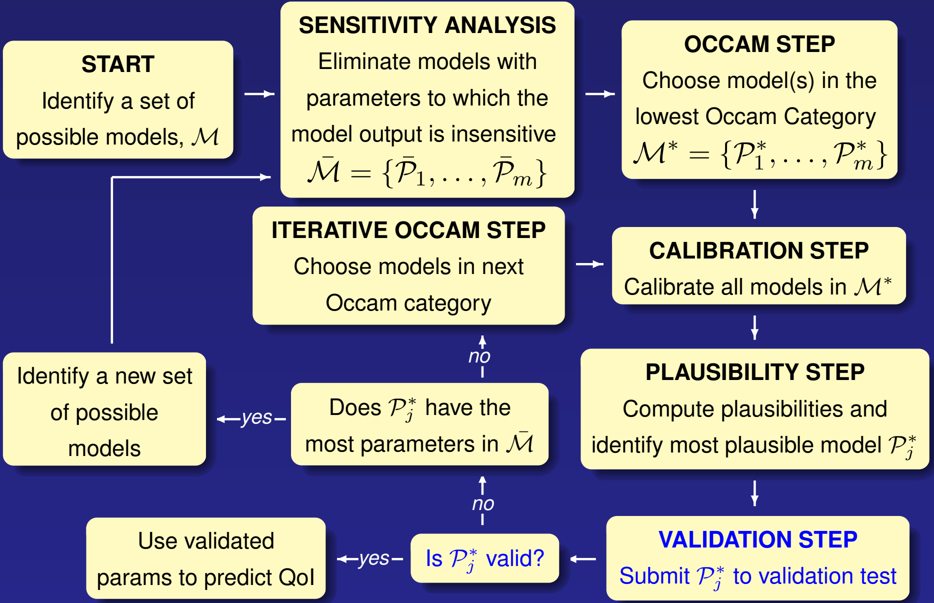
Here, $\mathbf{y}$ = potential energy of C_6H_{14}

Plausibility

$$\rho_j^* = \pi(\mathcal{P}_j^*|\mathbf{y},\mathcal{M}^*) = \frac{\pi(\mathbf{y}|\mathcal{P}_j^*,\mathcal{M}^*)\pi(\mathcal{P}_j^*|\mathcal{M}^*)}{\pi(\mathbf{y}|\mathcal{M}^*)}$$

Model	Bonds	Angles	Dihedrals	LJ 12-6	LJ 9-6	Params	Plausibility
$\mathcal{P}_1^*$	✓					2	1
$\mathcal{P}_2^*$		✓				2	0
$\mathcal{P}_3^*$				✓		2	0
$\mathcal{P}_4^*$					✓	2	0

The Occam-Plausibility Algorithm



Example: Occam-Plausibility Algorithm (cont)

As a validation scenario, we consider $C_{18}H_{38}$ at T = 300K in a canonical ensemble.

Validation

$$\pi(\boldsymbol{\theta}_1^*|\mathbf{y}_v, \mathbf{y}_c) = \frac{\pi(\mathbf{y}_v|\boldsymbol{\theta}_1^*, \mathbf{y}_c)\pi(\boldsymbol{\theta}_1^*|\mathbf{y}_c)}{\pi(\mathbf{y}_v)}$$

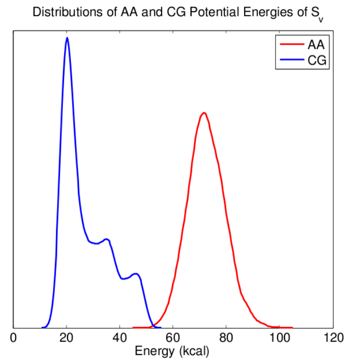
Here, $\mathbf{y}_v$ is the potential energy

How well does this updated model reproduce the desired observable?

Let

- $\pi(Q) = \pi(u_{AA}) \quad \Rightarrow \quad \pi(Q|\boldsymbol{\theta}^*) = \pi(U(\cdot;\boldsymbol{\theta}^*)), \;\; \gamma_{tol,1} = 0.05\sigma_{AA}^2$
- $Q = \langle u_{AA} \rangle \quad \Rightarrow \quad \mathbb{E}\left[\pi(Q|\boldsymbol{\theta}^*)\right] = \mathbb{E}\left[\langle U(\cdot;\boldsymbol{\theta}^*)\rangle\right], \; \gamma_{tol,2} = 0.2Q$

Example: Occam-Plausibility Algorithm (cont)



If we compare the distributions,

$$D_{KL}(\pi(Q_{AA})\|\pi(Q_{CG})) = 0.2435\sigma^2_{AA} > \gamma_{1,tol}$$

where $\gamma_{tol,1} = 0.05\sigma^2_{AA}$

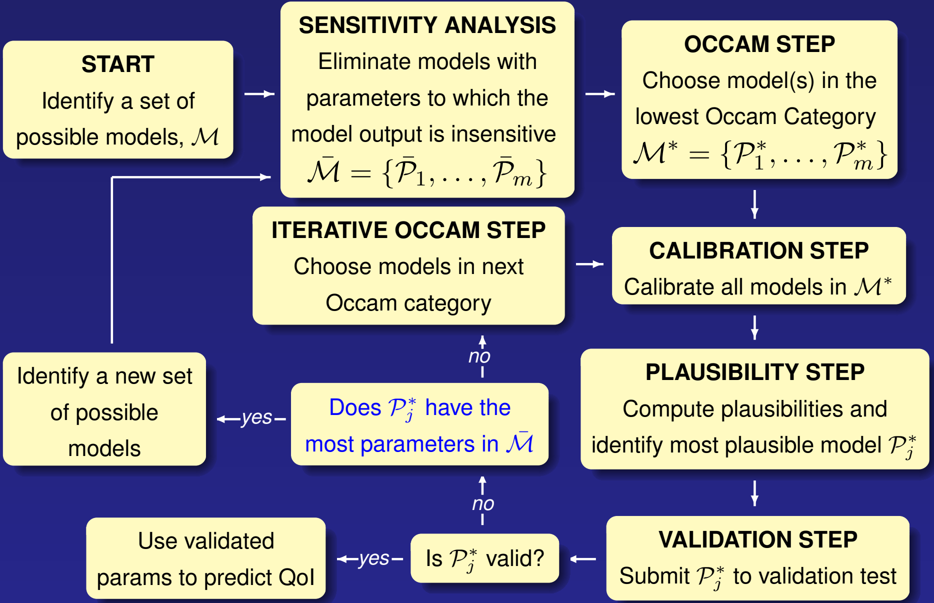
If we compare the ensemble average,

$$\left|Q_{AA} - \mathbb{E}_{\pi^v_{post}}[\pi(Q_v|\boldsymbol{\theta})]\right| = 0.67173Q > \gamma_{2,tol}$$

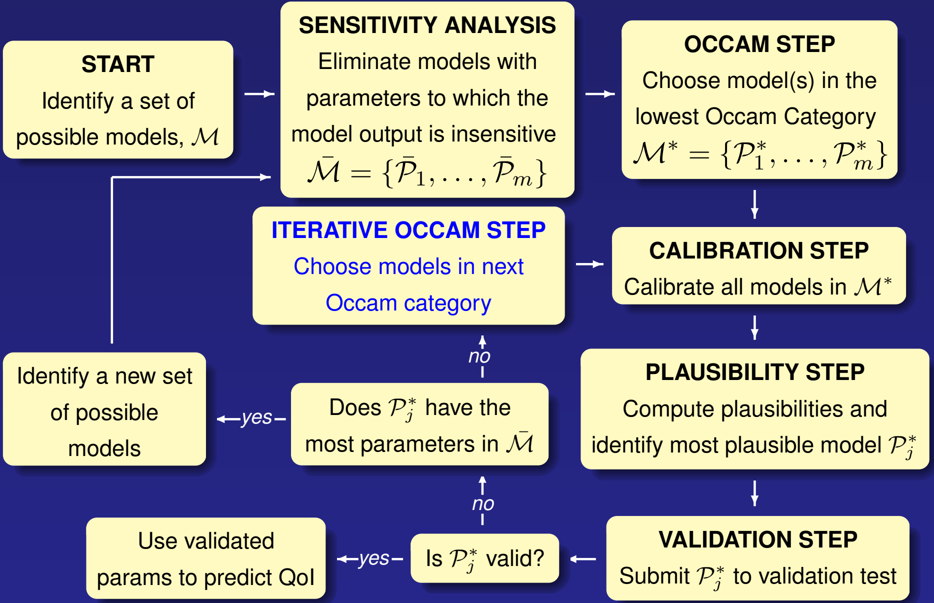
where $\gamma_{tol,2} = 0.2Q_{AA}$

Model is invalid

The Occam-Plausibility Algorithm



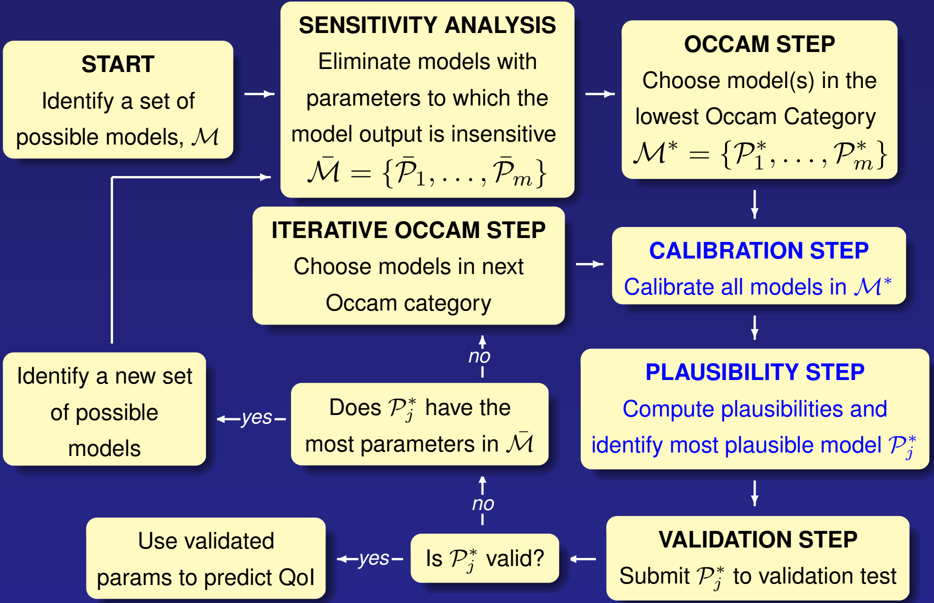
The Occam-Plausibility Algorithm



Example: Occam-Plausibility Algorithm (cont)

Model	Bonds	Angles	Dihedrals	LJ 12-6	LJ 9-6	Params	Category
$\bar{\mathcal{P}}_1$	✓					2	1
$\bar{\mathcal{P}}_2$		✓				2	
$\bar{\mathcal{P}}_3$				✓		2	
$\bar{\mathcal{P}}_4$					✓	2	
$\bar{\mathcal{P}}_5$	✓	✓				4	2
$\bar{\mathcal{P}}_6$	✓			✓		4	
$\bar{\mathcal{P}}_7$	✓				✓	4	
$\bar{\mathcal{P}}_8$		✓		✓		4	
$\bar{\mathcal{P}}_9$		✓			✓	4	
$\bar{\mathcal{P}}_{10}$	✓	✓		✓		6	3
$\bar{\mathcal{P}}_{11}$	✓			✓	✓	6	

The Occam-Plausibility Algorithm



Example: Occam-Plausibility Algorithm (cont)

Calibration

$$\pi(\boldsymbol{\theta}_j^*|\mathbf{y}, \mathcal{P}_j^*, \mathcal{M}^*) = \frac{\pi(\mathbf{y}|\boldsymbol{\theta}_j^*, \mathcal{P}_j^*, \mathcal{M}^*)\pi(\boldsymbol{\theta}_j^*|\mathcal{P}_j^*, \mathcal{M}^*)}{\pi(\mathbf{y}|\mathcal{P}_j^*, \mathcal{M}^*)}$$

Here, $\mathbf{y}$ = potential energy of C_6H_{14}

Plausibility

$$\rho_j^* = \pi(\mathcal{P}_j^*|\mathbf{y}, \mathcal{M}^*) = \frac{\pi(\mathbf{y}|\mathcal{P}_j^*, \mathcal{M}^*)\pi(\mathcal{P}_j^*|\mathcal{M}^*)}{\pi(\mathbf{y}|\mathcal{M}^*)}$$

Model	Bonds	Angles	Dihedrals	LJ 12-6	LJ 9-6	Params	Plausibility
$\mathcal{P}_1^*$	✓	✓				4	3.7891×10^{-7}
$\mathcal{P}_2^*$	✓			✓		4	0.3420
$\mathcal{P}_3^*$	✓				✓	4	0.6580

The Occam-Plausibility Algorithm

START
Identify a set of possible models, $\mathcal{M}$

SENSITIVITY ANALYSIS
Eliminate models with parameters to which the model output is insensitive
$\bar{\mathcal{M}} = \{\bar{\mathcal{P}}_1, \dots, \bar{\mathcal{P}}_m\}$

OCCAM STEP
Choose model(s) in the lowest Occam Category
$\mathcal{M}^* = \{\mathcal{P}_1^*, \dots, \mathcal{P}_m^*\}$

ITERATIVE OCCAM STEP
Choose models in next Occam category

CALIBRATION STEP
Calibrate all models in $\mathcal{M}^*$

Identify a new set of possible models

yes

Does $\mathcal{P}_j^*$ have the most parameters in $\bar{\mathcal{M}}$

no

PLAUSIBILITY STEP
Compute plausibilities and identify most plausible model $\mathcal{P}_j^*$

Use validated params to predict QoI

yes

Is $\mathcal{P}_j^*$ valid?

no

VALIDATION STEP
Submit $\mathcal{P}_j^*$ to validation test

Example: Occam-Plausibility Algorithm (cont)

As a validation scenario, we consider $C_{18}H_{38}$ at T = 300K in a canonical ensemble.

Validation

$$\pi(\boldsymbol{\theta}_3^*|\mathbf{y}_v, \mathbf{y}_c) = \frac{\pi(\mathbf{y}_v|\boldsymbol{\theta}_3^*, \mathbf{y}_c)\pi(\boldsymbol{\theta}_3^*|\mathbf{y}_c)}{\pi(\mathbf{y}_v)}$$

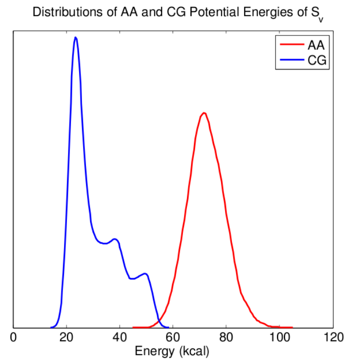
Here, $\mathbf{y}_v$ is the potential energy

How well does this updated model reproduce the desired observable?

Let

- $\pi(Q) = \pi(u_{AA}) \quad \Rightarrow \quad \pi(Q|\boldsymbol{\theta}^*) = \pi(U(\cdot;\boldsymbol{\theta}^*)), \;\; \gamma_{1,tol} = 0.05\sigma_{AA}^2$
- $Q = \langle u_{AA} \rangle \quad \Rightarrow \quad \mathbb{E}\left[\pi(Q|\boldsymbol{\theta}^*)\right] = \mathbb{E}\left[\langle U(\cdot;\boldsymbol{\theta}^*)\rangle\right], \;\; \gamma_{2,tol} = 0.2Q$

Example: Occam-Plausibility Algorithm (cont)



If we compare the distributions,

$$D_{KL}(\pi(Q_{AA}) \| \pi(Q_{CG})) = 0.2084\sigma^2_{AA} > \gamma_{1,tol}$$

where $\gamma_{tol,1} = 0.05\sigma^2_{AA}$

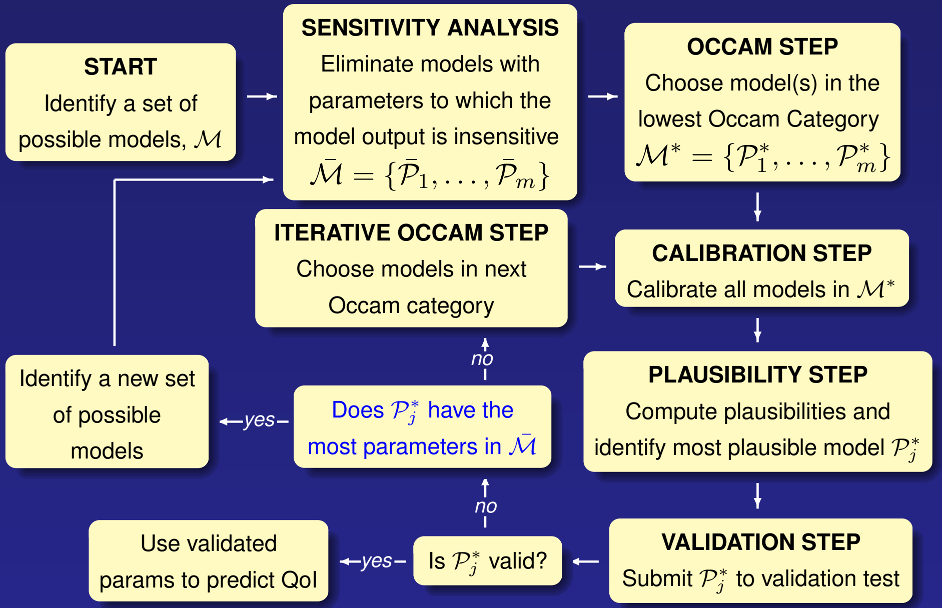
If we compare the ensemble average,

$$\left| Q_{AA} - \mathbb{E}_{\pi^v_{post}}[\pi(Q_v|\boldsymbol{\theta})] \right| = 0.5731Q > \gamma_{2,tol}$$

where $\gamma_{tol,2} = 0.2Q_{AA}$

Model is invalid

The Occam-Plausibility Algorithm



The Occam-Plausibility Algorithm

START

Identify a set of possible models, $\mathcal{M}$

SENSITIVITY ANALYSIS

Eliminate models with parameters to which the model output is insensitive

$\bar{\mathcal{M}} = \{\bar{\mathcal{P}}_1, \ldots, \bar{\mathcal{P}}_m\}$

OCCAM STEP

Choose model(s) in the lowest Occam Category

$\mathcal{M}^* = \{\mathcal{P}_1^*, \ldots, \mathcal{P}_m^*\}$

CALIBRATION STEP

Calibrate all models in $\mathcal{M}^*$

PLAUSIBILITY STEP

Compute plausibilities and identify most plausible model $\mathcal{P}_j^*$

VALIDATION STEP

Submit $\mathcal{P}_j^*$ to validation test

Is $\mathcal{P}_j^*$ valid?

yes → Use validated params to predict QoI

no → Does $\mathcal{P}_j^*$ have the most parameters in $\bar{\mathcal{M}}$

yes → Identify a new set of possible models

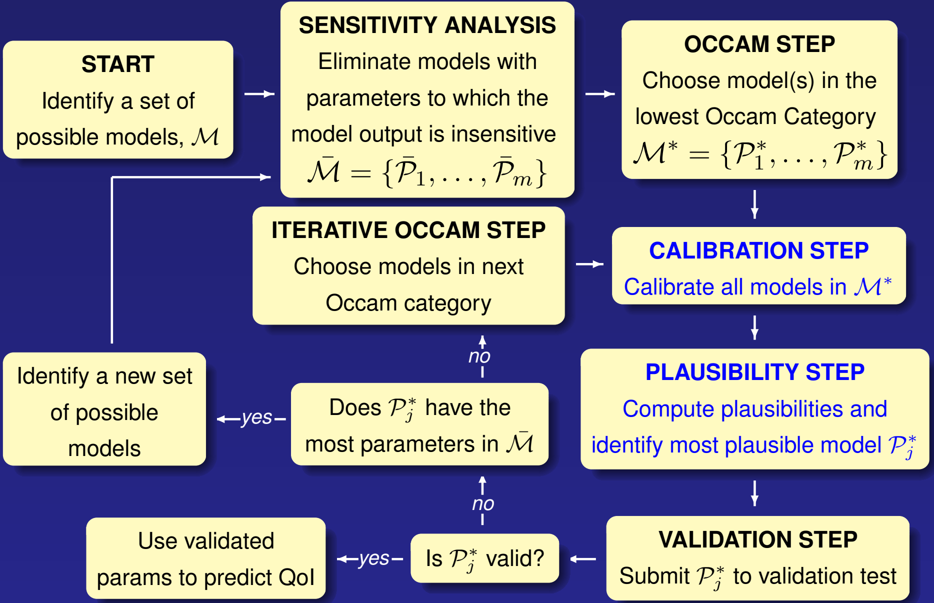
no → **ITERATIVE OCCAM STEP**

Choose models in next Occam category

Example: Occam-Plausibility Algorithm (cont)

Model	Bonds	Angles	Dihedrals	LJ 12-6	LJ 9-6	Params	Category
$\bar{\mathcal{P}}_1$	✓					2	1
$\bar{\mathcal{P}}_2$		✓				2	
$\bar{\mathcal{P}}_3$				✓		2	
$\bar{\mathcal{P}}_4$					✓	2	
$\bar{\mathcal{P}}_5$	✓	✓				4	2
$\bar{\mathcal{P}}_6$	✓			✓		4	
$\bar{\mathcal{P}}_7$	✓				✓	4	
$\bar{\mathcal{P}}_8$		✓		✓		4	
$\bar{\mathcal{P}}_9$		✓			✓	4	
$\bar{\mathcal{P}}_{10}$	✓	✓		✓		6	3
$\bar{\mathcal{P}}_{11}$	✓			✓	✓	6	

The Occam-Plausibility Algorithm



Example: Occam-Plausibility Algorithm (cont)

Calibration

$$\pi(\boldsymbol{\theta}_j^*|\mathbf{y},\mathcal{P}_j^*,\mathcal{M}^*) = \frac{\pi(\mathbf{y}|\boldsymbol{\theta}_j^*,\mathcal{P}_j^*,\mathcal{M}^*)\pi(\boldsymbol{\theta}_j^*|\mathcal{P}_j^*,\mathcal{M}^*)}{\pi(\mathbf{y}|\mathcal{P}_j^*,\mathcal{M}^*)}$$

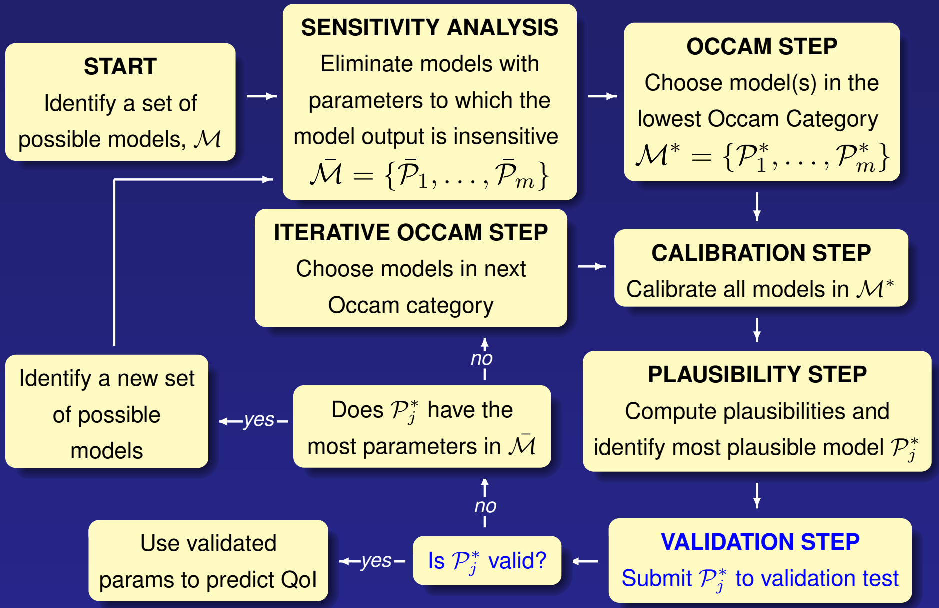
Here, $\mathbf{y}$ = potential energy of C_6H_{14}

Plausibility

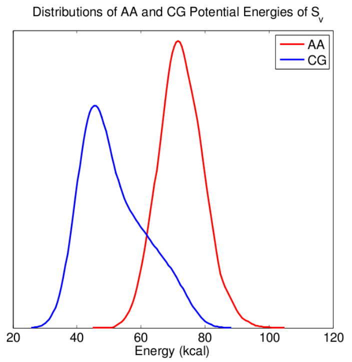
$$\rho_j^* = \pi(\mathcal{P}_j^*|\mathbf{y},\mathcal{M}^*) = \frac{\pi(\mathbf{y}|\mathcal{P}_j^*,\mathcal{M}^*)\pi(\mathcal{P}_j^*|\mathcal{M}^*)}{\pi(\mathbf{y}|\mathcal{M}^*)}$$

Model	Bonds	Angles	Dihedrals	LJ 12-6	LJ 9-6	Params	Plausibility
$\bar{\mathcal{P}}_{10}$	✓	✓		✓		6	0.5
$\bar{\mathcal{P}}_{11}$	✓			✓	✓	6	0.5

The Occam-Plausibility Algorithm



Example: Occam-Plausibility Algorithm (cont)



If we compare the distributions,

$$D_{KL}(\pi(Q_{AA})\|\pi(Q_{CG})) = 0.0452\sigma_{AA}^2 < \gamma_{1,tol}$$

where $\gamma_{tol,1} = 0.05\sigma_{AA}^2$

If we compare the ensemble average,

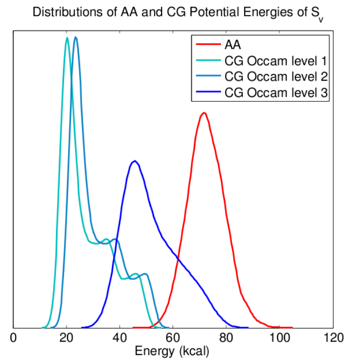
$$\left|Q_{AA} - \mathbb{E}_{\pi_{post}^v}[\pi(Q_v|\boldsymbol{\theta})]\right| = 0.1721Q < \gamma_{2,tol}$$

where $\gamma_{tol,2} = 0.2Q_{AA}$

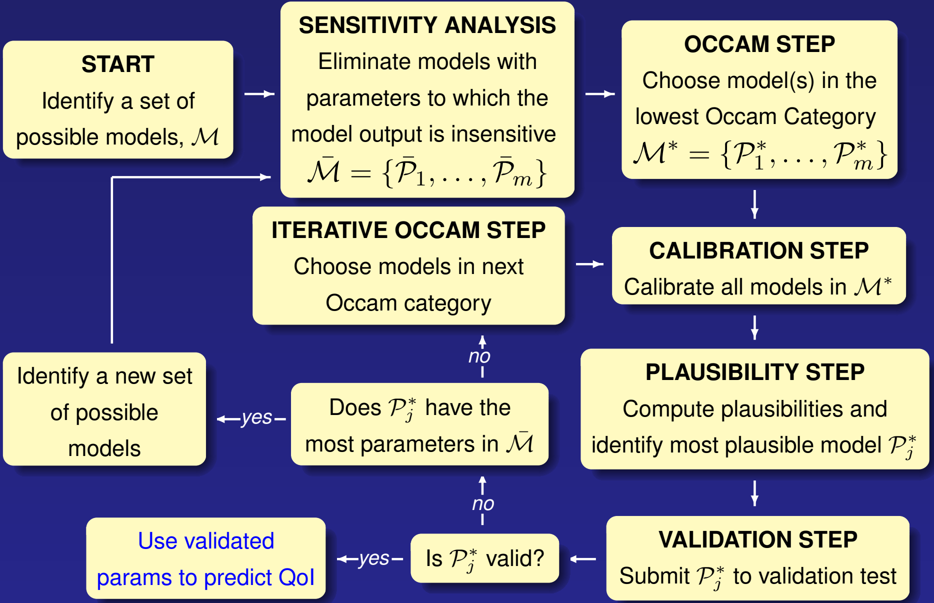
Model is NOT invalid

Example: Occam-Plausibility Algorithm (cont)

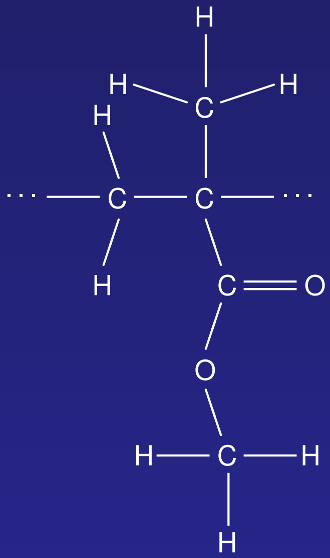
How do the observables change as we move through the Iterative Occam Step?



The Occam-Plausibility Algorithm

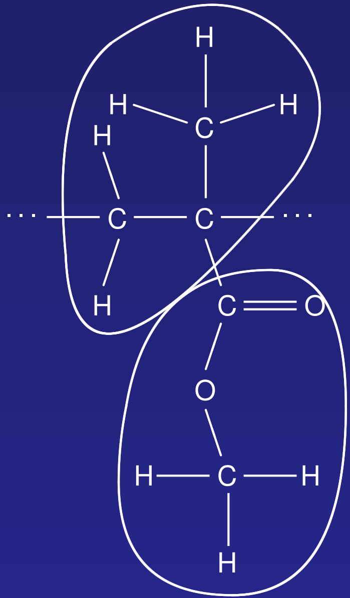


Work in Progress: PMMA

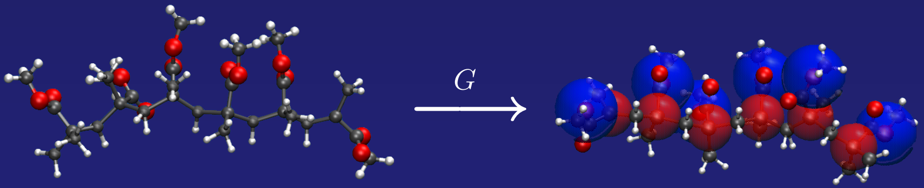


One molecule has:
15 atoms
72 parameters

Work in Progress: PMMA

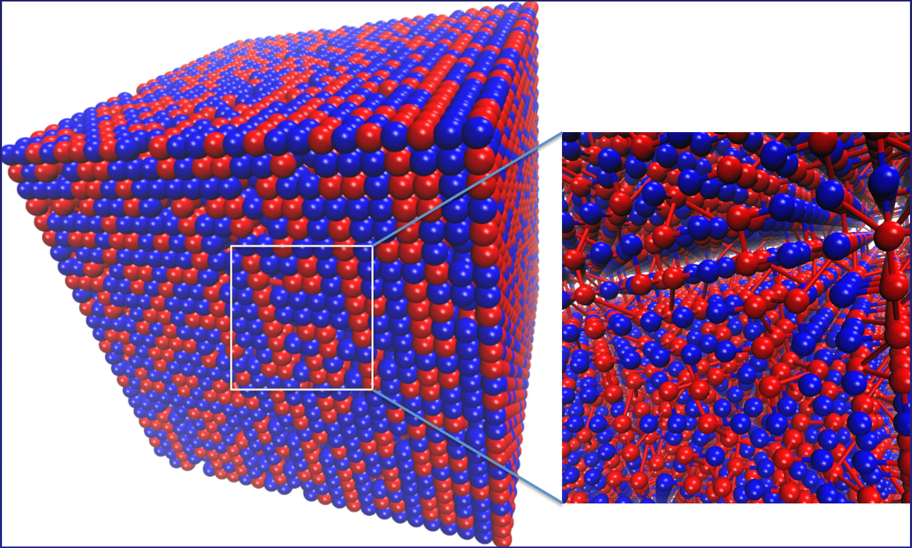


CG Calibration Scenario



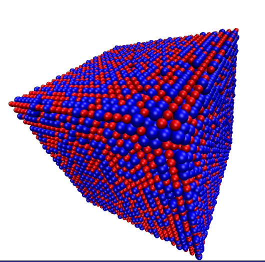
Model	Bonds	Angles	LJ 9-6	LJ 12-6	A	# of Parameters
$\mathcal{P}_1$	✓			✓	✓	9
$\mathcal{P}_2$	rigid	✓		✓	✓	9
$\mathcal{P}_3$	✓	✓		✓	✓	13
$\mathcal{P}_4$	✓		✓		✓	9
$\mathcal{P}_5$	rigid	✓	✓		✓	9
$\mathcal{P}_6$	✓	✓	✓		✓	13
$\mathcal{P}_7$	✓				✓	5
$\mathcal{P}_8$	rigid	✓			✓	5
$\mathcal{P}_9$	✓	✓			✓	9

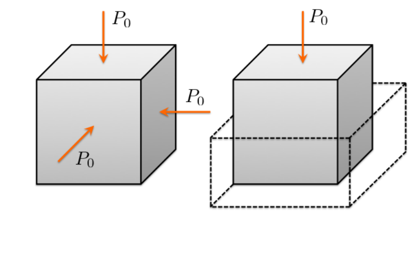
The polymerization process (KMC)



10x10x10 nm

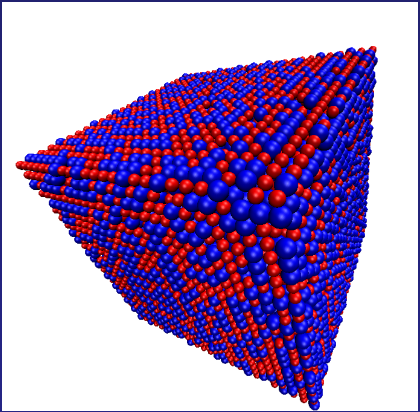
Continuum Models Calibration Scenario



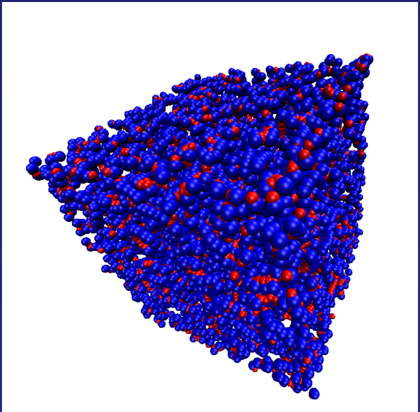


Model	# of Parameters
$\mathcal{P}_1$: Saint Venant-Kirchhoff	2
$\mathcal{P}_2$: Neo-Hookean	2
$\mathcal{P}_3$: Mooney-Rivilin	3

Continuum Models Calibration Scenario



Initial Configuration



Equlibration Configuration

Continuum Models Calibration Scenario

(PMMA Lattice under biaxial deformation)

Continuum Models Calibration Scenario

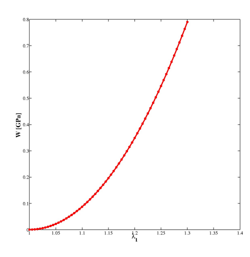
Hyperelastic Model

$\mathcal{P}_2$: **Compressible Neo-Hookean model**

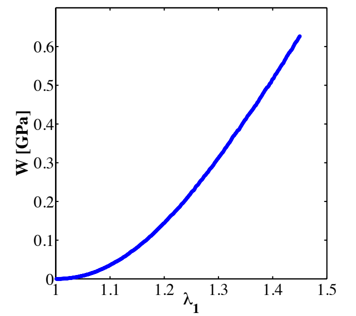
$$W = C_1(I_1 - 3) - 2C_1 \ln \sqrt{I_3} + C_2(\sqrt{I_3} - 1)^2$$

macromodel parameters $= \boldsymbol{\theta}_2 = (C_1, C_2)$

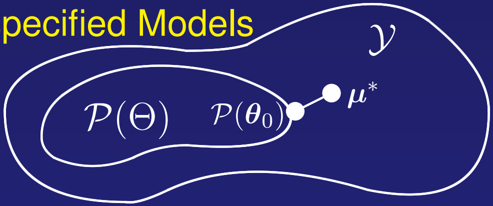
Biaxial deformation $\lambda_1 = \lambda_2$:
Strain-Energy $\rightarrow W = W(\lambda_1, \boldsymbol{\theta}_2)$



Observational data supplied by CG model to calibrate the continuum models



6. Model Inadequacy - Misspecified Models



Suppose $\mu^* \notin \mathcal{P}(\Theta)$. Then the density $g(\mathbf{y}) \notin \mathcal{P}(\Theta)$ (model is inadequate)

Let there exist a $\boldsymbol{\theta}_0$ such that

$$\boldsymbol{\theta}_0 = \underset{\boldsymbol{\theta}\in\Theta}{\operatorname{argmin}}\, D_{KL}(g(\mathbf{y})\|\pi(\cdot|\boldsymbol{\theta})) \quad \forall i$$

Then, under suitable smoothness conditions

$$\left\| \pi^n(y_1, y_2, \dots, y_n|\cdot) - \mathcal{N}\left(\hat{\boldsymbol{\theta}}_n, \frac{1}{n} V_{\boldsymbol{\theta}_0}\right) \right\|_{TV} \to 0 \quad as\ n \to \infty$$

where $\hat{\boldsymbol{\theta}}_n =$ the MLE and

$$(V_{\boldsymbol{\theta}_0})_{ij} = -\mathbb{E}_g\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j} L_{\mathbf{y}}(\boldsymbol{\theta}_0)\right]$$

Kleijn and van der Vaart (2012), Freedman, D. (2006), Geyer (2003)

Lemma: θ_0 is the maximum likelihood estimate

Proof:

$$\begin{aligned}\boldsymbol{\theta}_0 &= \underset{\Theta}{\operatorname{argmin}} \left[\int_{\mathcal{Y}} \left(g(\mathbf{y}) \log g(\mathbf{y}) - g(\mathbf{y}) \pi(\mathbf{y} \| \boldsymbol{\theta}) \right) \, d\mathbf{y} \right] \\ &= \underset{\Theta}{\operatorname{argmin}} \left[- \int_{\mathcal{Y}} g(\mathbf{y}) \pi(\mathbf{y} \| \boldsymbol{\theta}) \, d\mathbf{y} \right] \\ &= \underset{\Theta}{\operatorname{argmax}} \left[\int_{\mathcal{Y}} g(\mathbf{y}) \pi(\mathbf{y} \| \boldsymbol{\theta}) \, d\mathbf{y} \right] \\ &= \underset{\Theta}{\operatorname{argmax}} \, \mathbb{E}_g \left[\log \pi(\mathbf{y} | \boldsymbol{\theta}) \right]\end{aligned}$$

Model Misspecification and Model Plausibility

Theorem 1: (Bayesian $\rightarrow$ Frequentist)

If

$$\pi(\mathbf{y}|\mathcal{P}_i, \mathcal{M}) = \int_{\Theta} \pi(\mathbf{y}|\boldsymbol{\theta}, \mathcal{P}_i, \mathcal{M})\delta(\boldsymbol{\theta} - \boldsymbol{\theta}_0)\, d\boldsymbol{\theta} = \pi(\mathbf{y}|\boldsymbol{\theta}_0, \mathcal{P}_i, \mathcal{M})$$

and

$$\frac{\rho_1}{\rho_2} = \frac{\pi(\mathbf{y}|\boldsymbol{\theta}_{0,1}, \mathcal{P}_1, \mathcal{M})}{\pi(\mathbf{y}|\boldsymbol{\theta}_{0,2}, \mathcal{P}_2, \mathcal{M})} \times O_{12}$$

Then, if $\mathcal{P}_1$ is more plausible than $\mathcal{P}_2$ and $O_{12} \leq 1$,

$$D_{KL}(g \| \pi(\mathbf{y}|\boldsymbol{\theta}_{0,1}, \mathcal{P}_1, \mathcal{M})) < D_{KL}(g \| \pi(\mathbf{y}|\boldsymbol{\theta}_{0,2}, \mathcal{P}_2, \mathcal{M}))$$

(The converse: $D_{KL}(1) < D_{KL}(2) \Rightarrow \rho_1 > \rho_2$, holds only under special assumptions)

Model Misspecification and Model Plausibility

Corollary

For given observational data **y** , let $\mathcal{P}_1(\boldsymbol{\theta}_1)$ be the only well specified model in a set $\mathcal{M}$ of parametric models $\{\mathcal{P}_1(\boldsymbol{\theta}_1), \mathcal{P}_2(\boldsymbol{\theta}_2), \ldots, \mathcal{P}_m(\boldsymbol{\theta}_m)\}$ Then,
a) $\mathcal{P}_1(\boldsymbol{\theta}_1)$ is the most plausible model in the set $\mathcal{M}$,

$$\rho_1 > \rho_k, \quad k = 2, 3, \ldots, m,$$

b) there exists $\boldsymbol{\theta}^*$ belonging to $\mathcal{P}_1(\Theta_1)$ such that

$$\boldsymbol{\theta}^* = \mathrm{argmin} D_{KL}(g \| \pi(\mathbf{y}|\boldsymbol{\theta}))$$

Conclusions

- Bayes' theorem provides a powerful framework for dealing with model validation and uncertainty quantification
- The test of model validity can involve a sequence of statistical inverse problems for model parameters, each reflecting the projected influence of the QoI

> Validation is the process of determining the level of confidence one has in the ability of the model to predict quantities of interest based on the accuracy with which the model predicts specific observables to within preset tolerances

- The concept of model plausibility provides a powerful tool for
  1. determining potentials for CG models of atomic systems
  2. choosing models among a class of models that have parameter closest "in D_{KL}" to the true distribution

Conclusions

- Model inadequacy can be attributed to model misspecification: $\boldsymbol{\theta}^* \notin \mathcal{P}(\Theta)$
- The calculation of model sensitivities due to variations in parameters can significantly reduce the number of relevant models for given outputs.
- Hierarchical categories of models based on numbers of parameters (the Occam categories) together with the evaluation of model plausibilities provide a basis for an adaptive process for validation of parametric classes of coarse-grained models

Thank you!