A methodology for assessing uncertainties in simulation predictions

Kenneth M. Hanson

– – – – – – – – – – –

CCS-2, Methods for Advanced Scientific Simulation
Computer and Computational Science Division
Los Alamos National Laboratory

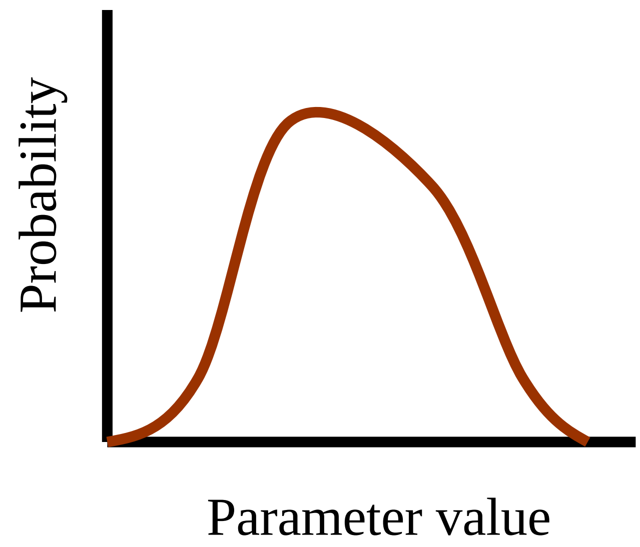
Presentation available under http://home.lanl.gov/kmh/

Overview

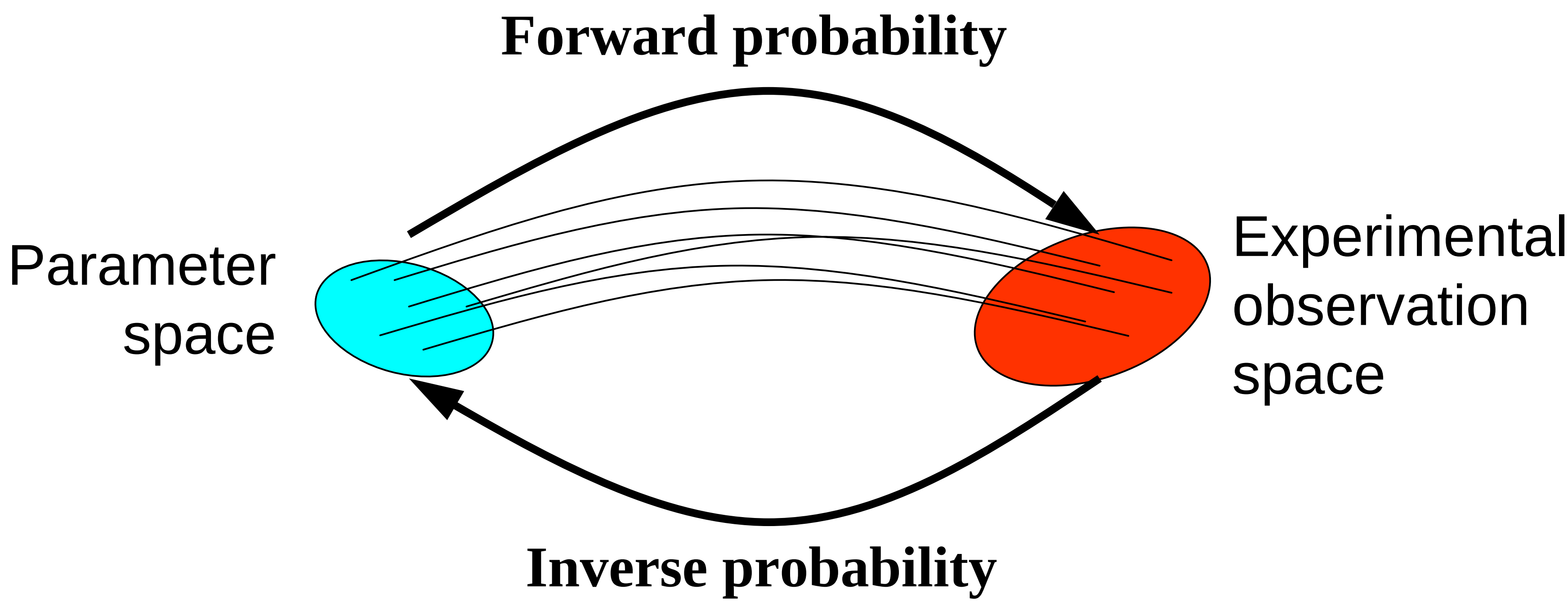
- Uncertainties
  - represented by probabilities
  - use Monte Carlo to visualize and estimate uncertainties
- Example - Taylor impact test
- Approach to validation of simulation code
  - focus is on uncertainties in simulation predictions
  - conduct validation experiments at various levels of integration of pertinent effects
- Ultimate goal: develop models that are consistent with all experiments

Uncertainty analysis

- Uncertainties in parameters are characterized by probability density functions (pdf)
- Probability interpreted as quantitative measure of "degree of belief"
- Rules of classical probability theory apply
- Bayes law provides way to update knowledge about models as summarized in terms of uncertainty



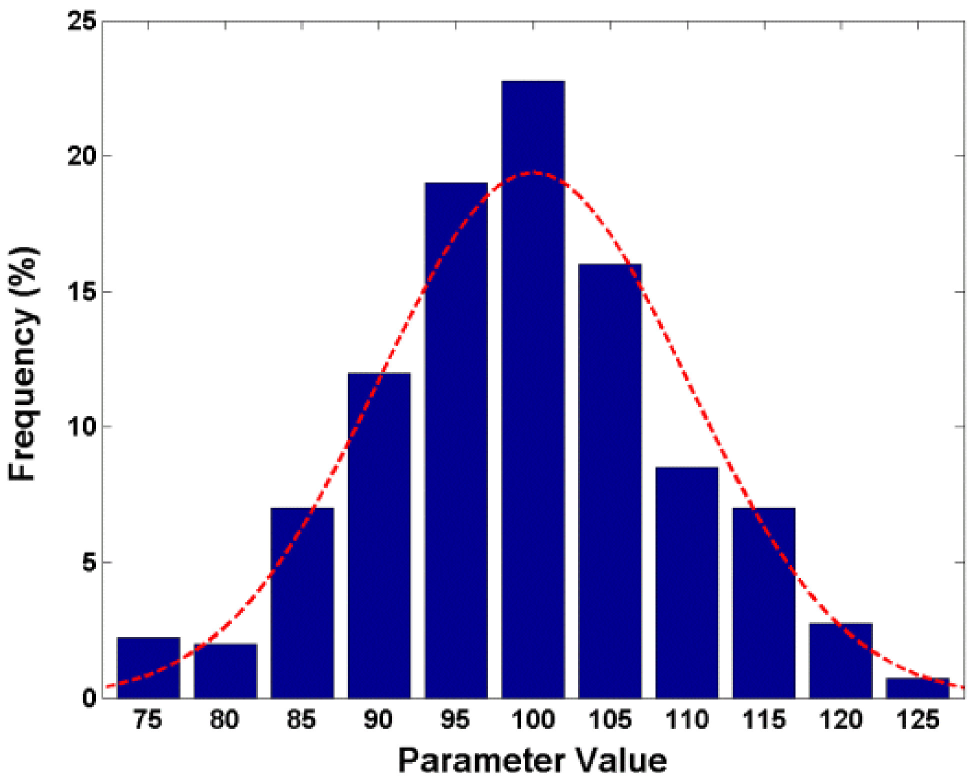
Forward and inverse probability



- Forward probability - determine uncertainties in observables resulting from model parameter uncertainties
- Inverse probability - infer model parameters and their uncertainties from uncertainties in observables

Monte Carlo technique

- Monte Carlo -
  - ▸ numerical technique to do probabilistic calculations
  - ▸ draw values from prob. density function (pdf)
  - ▸ use these values in numerical calculation
- Figure shows histogram of 100 parameter values randomly drawn from Gaussian pdf



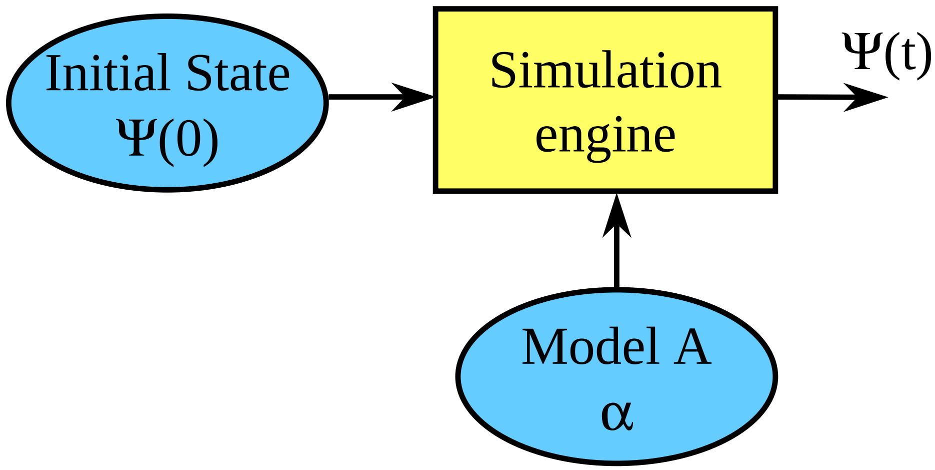
Monte Carlo technique

- Represent probability density function by a set of numbers drawn randomly from it
  - consider a parameter space of n dimensions represented by vector $\mathbf{x}$
  - probability density function (pdf), $p(\mathbf{x})$
  - draw a sequence of random samples $\{\mathbf{x}_k\}$ from it
- Allows evaluation of expectation values
  - for K samples,

$$\langle f(\mathbf{x}) \rangle = \int f(\mathbf{x})\, q(\mathbf{x})\, d\mathbf{x} \cong (1/K)\, \Sigma_k\, f(\mathbf{x}_k)$$

  - typical use is to calculate mean $\langle \mathbf{x} \rangle$ and variance $\langle (\mathbf{x} - \langle \mathbf{x} \rangle)^2 \rangle$ of pdf

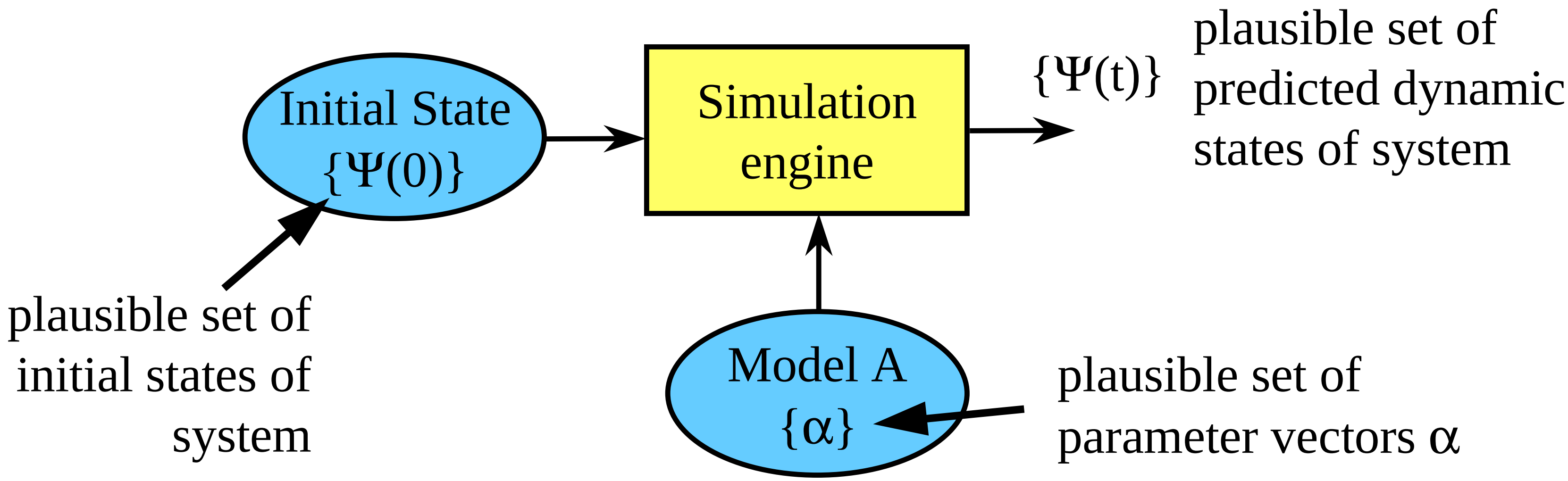
Schematic view of simulation code



- Simulation code predicts state of time-evolving system

  $\Psi(t)$ = time-dependent state of system
- Requires as input
  - $\Psi(0)$ = initial state of system
  - description of physics behavior of each system component; e.g., physics model A with parameter vector α (e.g., constitutive relations)
- Simulation engine solves the dynamical equations (PDEs)

Simulation of plausible simulation predictions



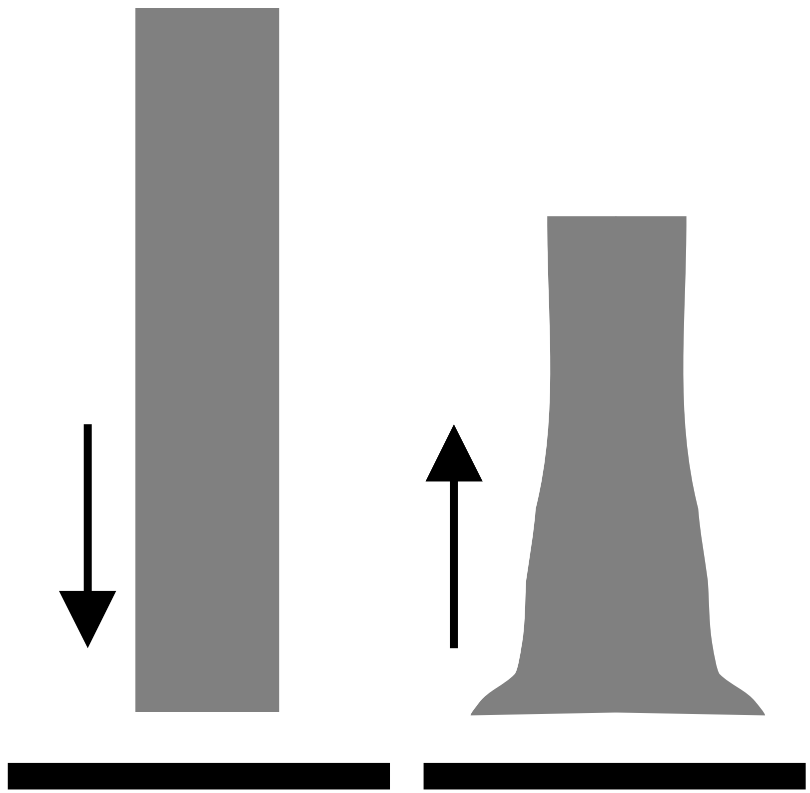
- Generate plausible predictions for known uncertainties in parameters and initial conditions
- Monte Carlo method
  - run simulation code for each random draw from pdf for α, $p(\alpha|.)$, and initial state, $p(\Psi(0)\,|.)$
  - simulation outputs represent plausible set of predictions, $\{\Psi(t)\}$

Uncertainty in predictions

- Estimate by propagating through simulation code parameter samples drawn from joint posterior distribution of all parameters describing constituent physics models
- Assumptions about simulation code:
  - appropriate physics modules included
  - simulation uncertainties dominated by uncertainties in physics modules, which can be determined through carefully designed experiments (validation issue)
  - numerically accurate (verification issue)
- Other stochastic effects may be included
  - variability in material properties, e.g., densities, grain structure
  - chaotic behavior

Taylor impact test

- Propel cylinder into rigid flat plate
- Measure profile of deformed cylinder
- Deformation depends on
  - cylinder dimensions
  - impact velocity
  - plastic flow behavior of material at high strain rate
- Useful for
  - determining parameters in material-flow model
  - validating simulation code (including material model)

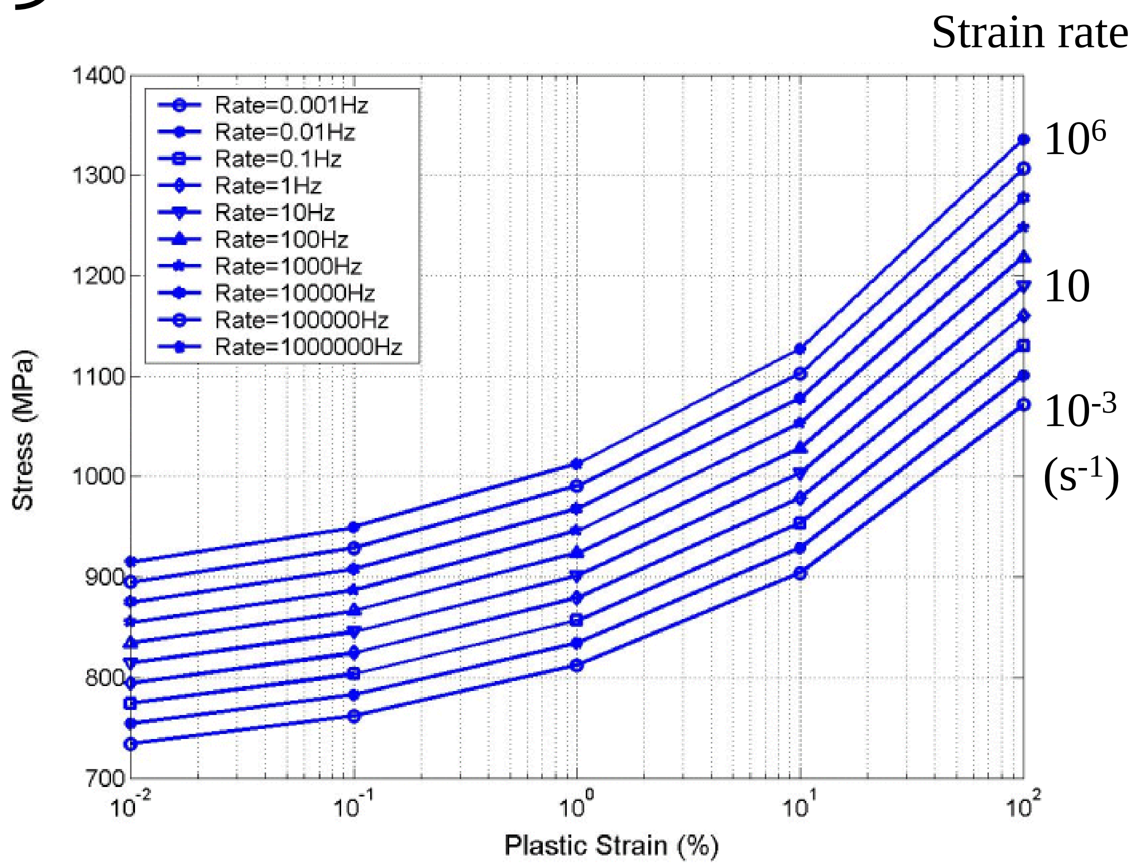


Stress-strain relation

- Johnson-Cook model for rate-dependent strength and plasticity

$$\sigma = (\alpha_1 + \alpha_2 \varepsilon_p^N)\left[1 + \alpha_3 \log\left(\frac{\partial \varepsilon_p}{\partial t}\right)\right]$$

- Parameters can be determined from Hopkinson bar measurements



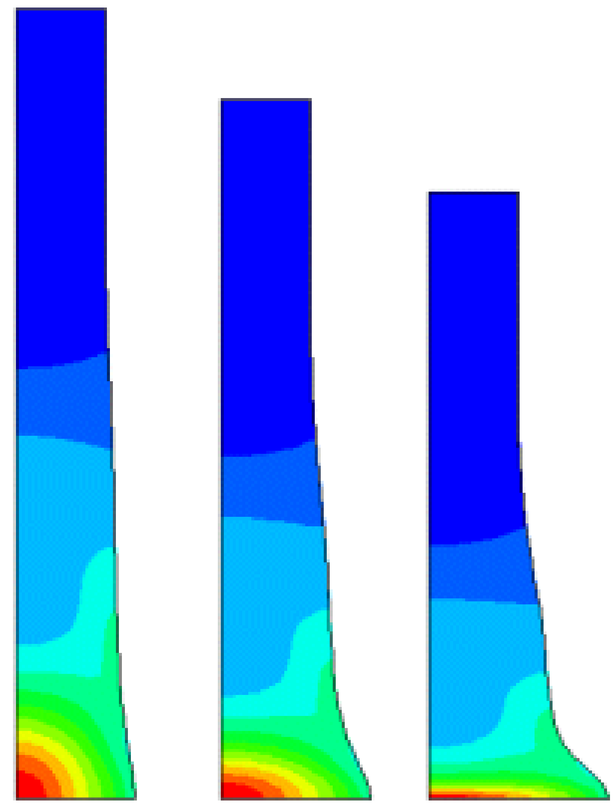
Taylor test simulations

- Simulate Taylor impact test
  - Abaqus FEM code
  - Johnson-Cook model for rate-dependent strength and plasticity

$$\sigma = (\alpha_1 + \alpha_2 \varepsilon_p^N)\left[1 + \alpha_3 \log\left(\frac{\partial \varepsilon_p}{\partial t}\right)\right]$$

  (4 parameters)

  - ignore anisotropy, fracture effects
  - cylinder: high-strength steel
    15-mm dia, 38-mm long
  - impact velocity = 350 m/s
- Effective strain reaches 250%

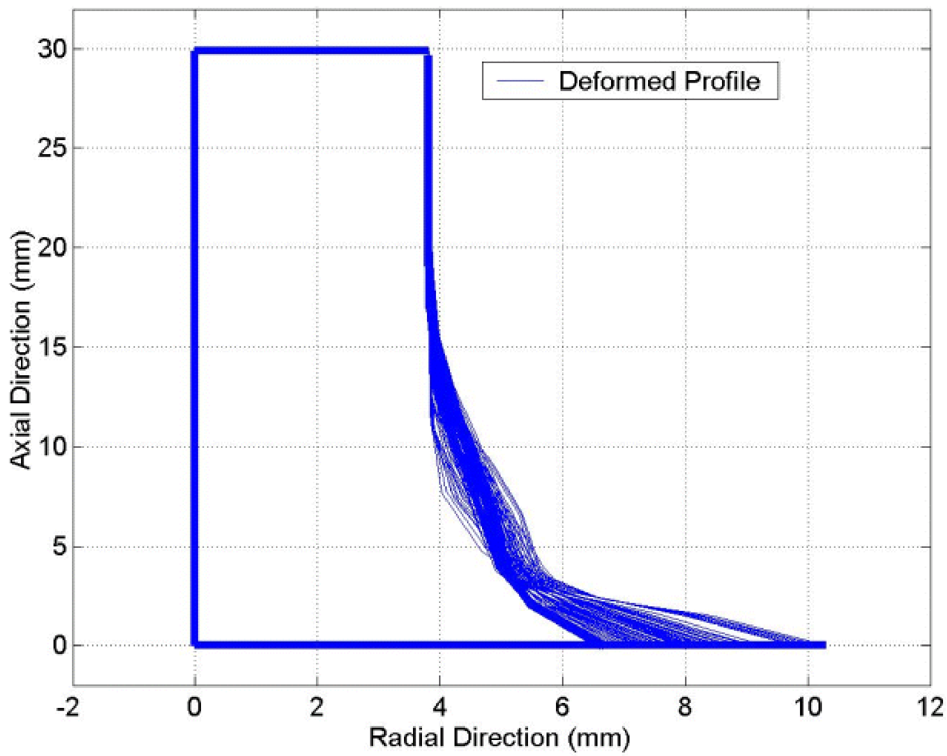


17 μs 33 μs 50 μs

Monte Carlo example - Taylor test

- Use MC technique to propagate uncertainties through deterministic simulation code
    - Assume uncertainty in each parameter in Johnson-Cook model (20-40%)
    - Draw value for each of four parameters from its assumed Gaussian pdf
    - Run Abaqus code for each set of parameters
- Figure shows range of variation in predicted cylinder shape

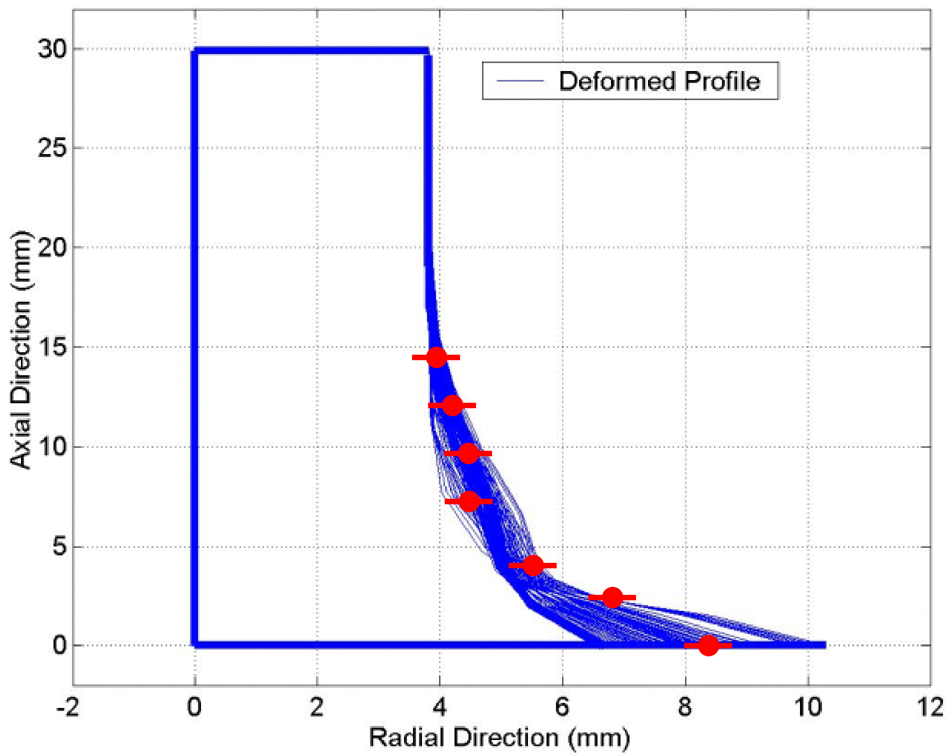
NESSUS/Abaqus



Comparison with experiment

- Suppose we do an experiment that replicates the conditions of the simulation and measure the profile after impact
- Quantitative comparison of simulation prediction with experimental data must take into account uncertainties in both

NESSUS/Abaqus



Visualization of uncertainty

- Problem inherently difficult for numerous variables, especially for fields, e.g., stress or strain vs. (x, y, z, t)
- With Monte Carlo, use normal tools to visualize simulation
  - view several plausible realizations from MC sequence
  - view MC sequence as video loop (for field at fixed time)

Inference

- May want to make inference about quantities that have not been directly measured
  - stress
  - pressure
  - temperature

- Use validation experiment to update info about model
  - capture info in terms of uncertainties
  - uncertainties indicate degree of confidence in prediction
  - attempt to develop model that is consistent with ALL available experiments Inference - unmeasured quantities, new situations, conditions

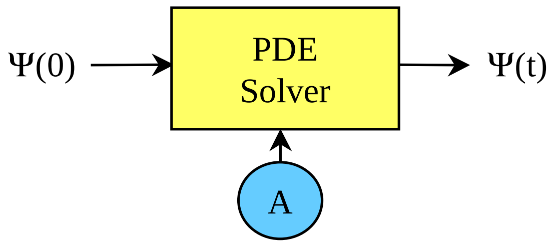
Issues

- Hierarchy of experiments
  - basic - designed to isolate and characterize a basic physical phenomenon at single
  - partially integrated - involves more complex combination of phenomena, e.g., multiple materials, varying conditions, complex geometry, ...
  - fully integrated - attempt to approach application conditions
- inference - use validation experiment to update info about model
  - capture info in terms of uncertainties
  - uncertainties indicate degree of confidence in prediction
  - attempt to develop model that is consistent with ALL available experiments
- Ultimate goal Combine results from many (all) experiments
  - reduce uncertainties in model parameters
  - require consistency of models with all experiments

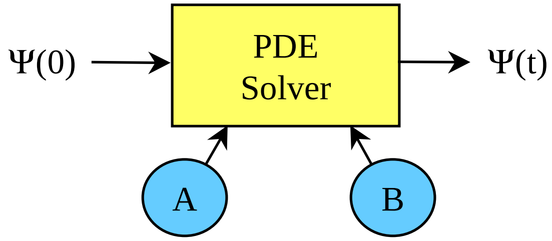
Validation Experiments

Full validation requires hierarchy of experiments

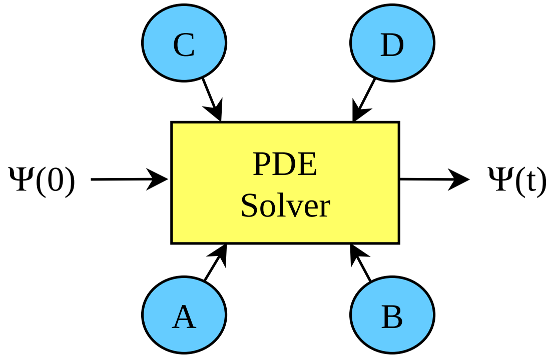
- **Basic** experiments determine individual physics models



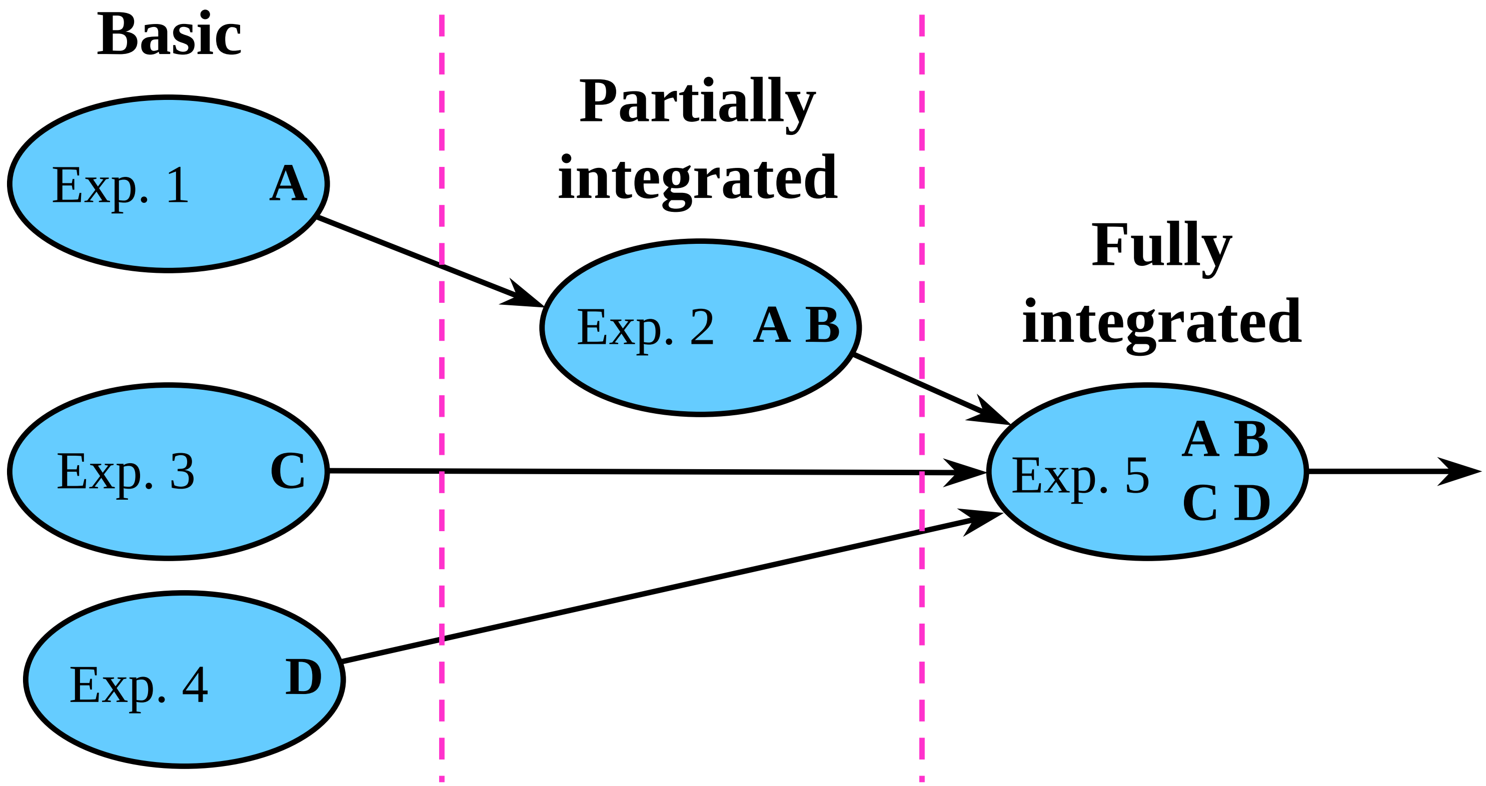
- **Partially integrated** experiments involve combinations of two or more elemental models



- **Fully integrated** experiments require complete set of models needed to describe final application of simulation code



Hierarchy of experiments

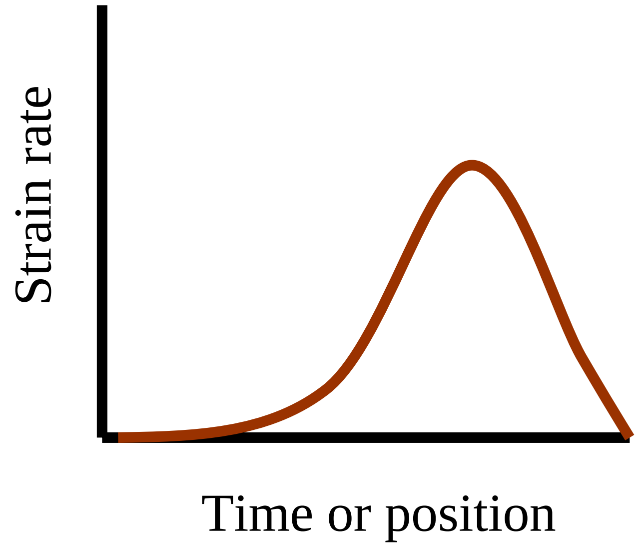


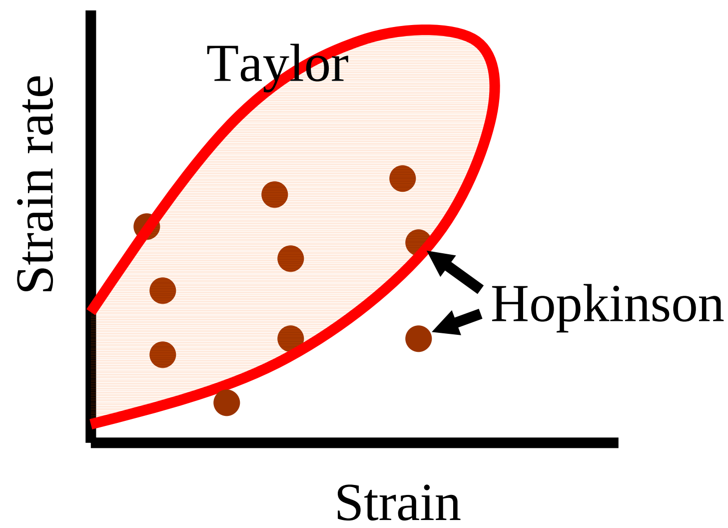
- Information flow in analysis of series of experiments
- Information about models accumulates

Hierarchy of experiments - plasticity

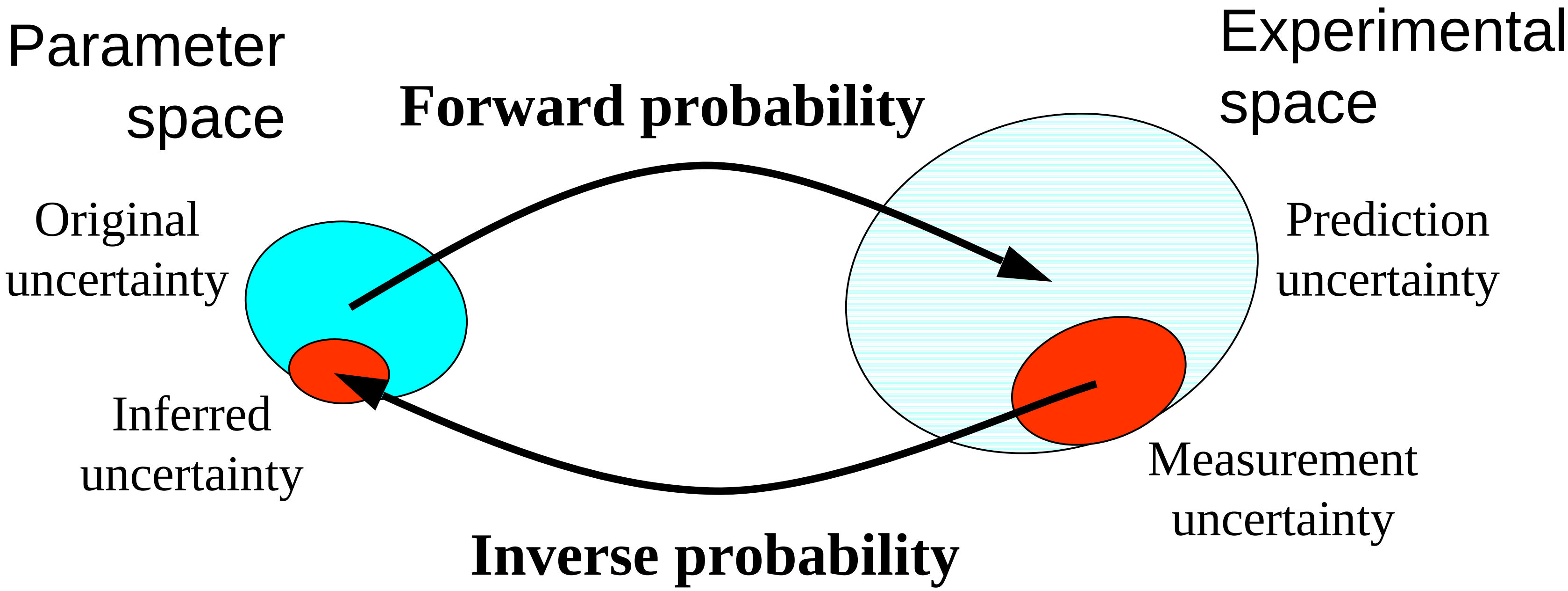
- Hierarchy of experiments
  - basic - Hopkinson bar
    measures stress-strain relationship at specific stain and strain rate
  - partially integrated - Taylor test
    covers range of strain rates

    may extend range of physical conditions
  - full integrated - application dependent
    - pressure vessel
    - bumper rail
    - automobile crash test



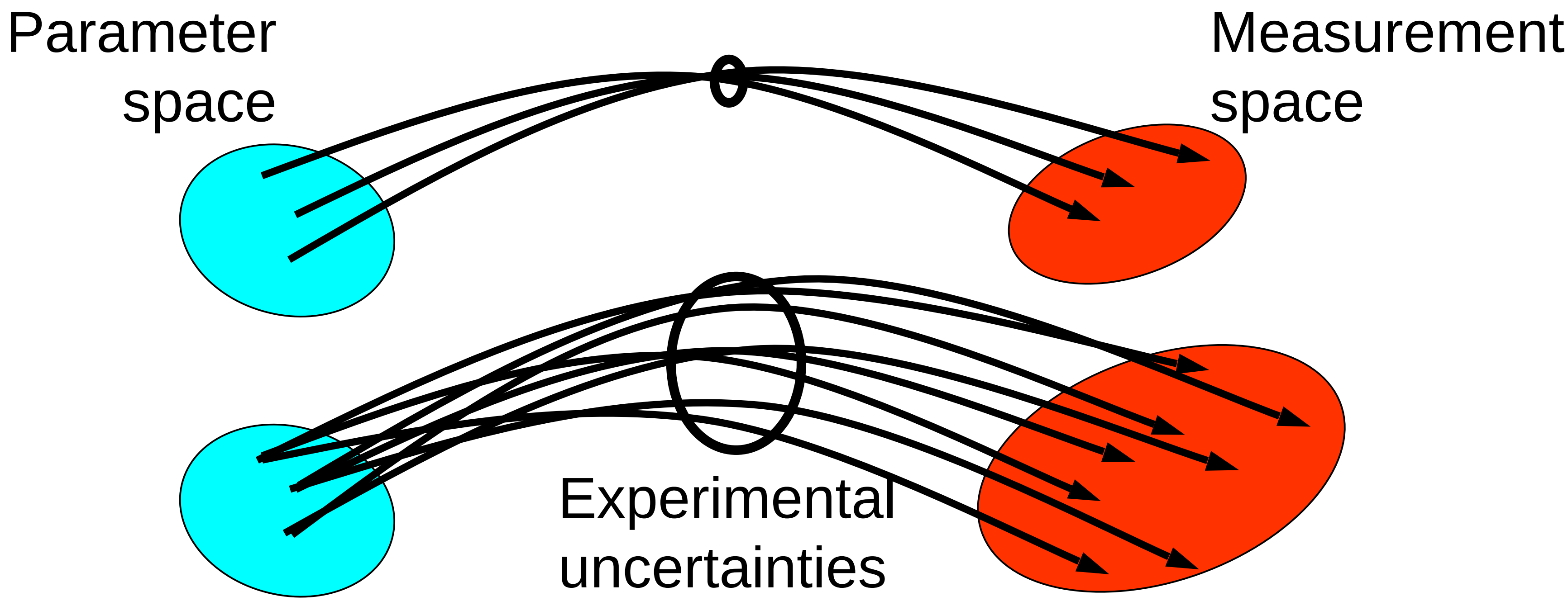


Forward and inverse probability



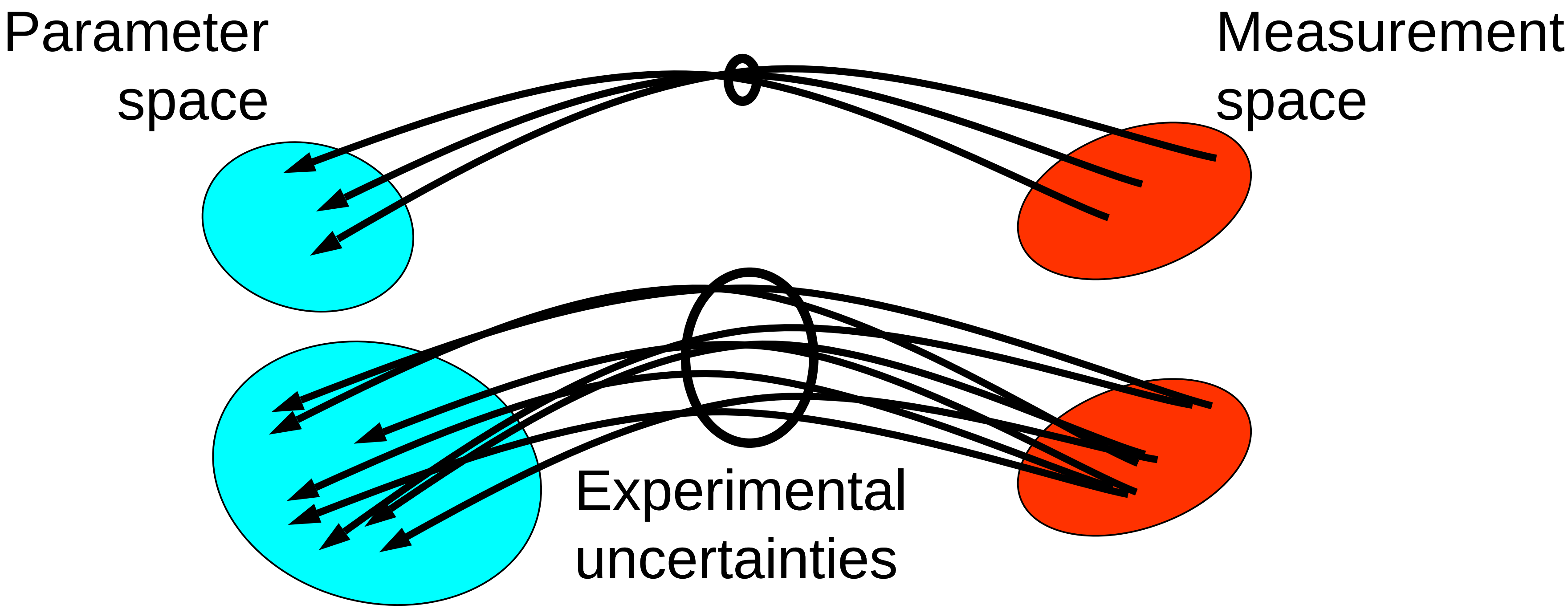
- Model inference
  - if uncertainties in measurements are smaller than prediction uncertainties in that arise from parameter uncertainties,
    one may be able to reduce uncertainties in parameters
  - conditional on prediction uncertainties being dominated by uncertainties in parameters and not by those in experimental set up
  - highlights importance of **good experimental technique**

Forward probability



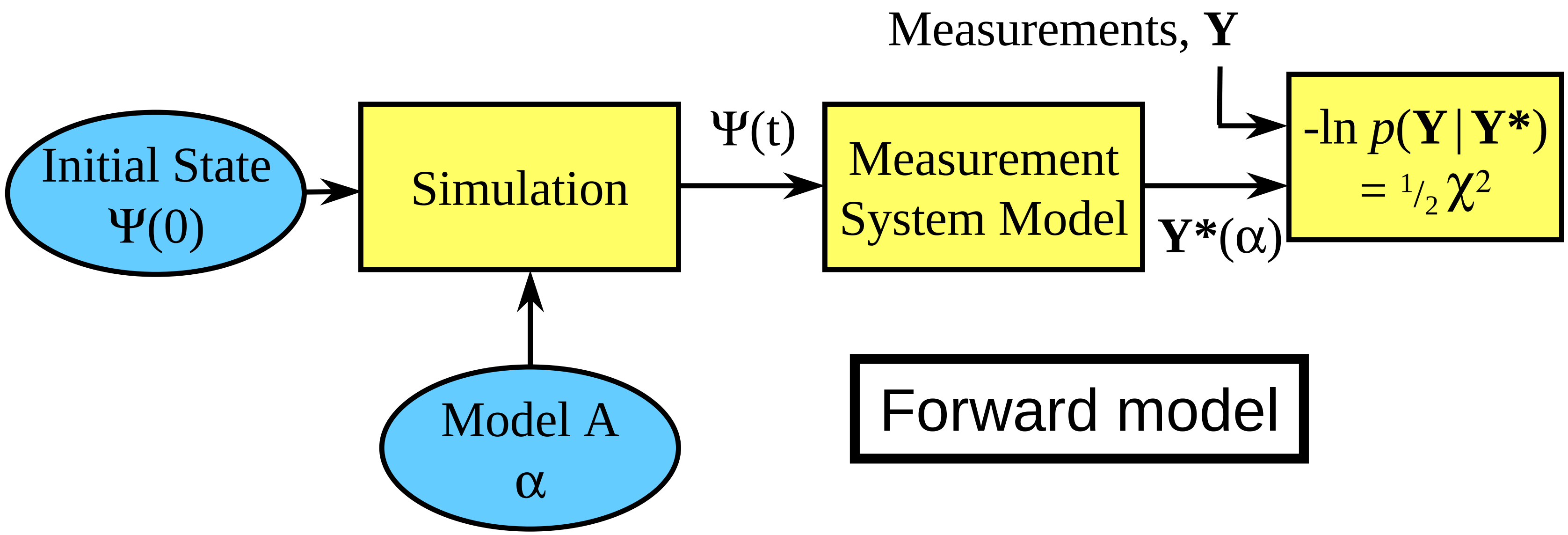
- Uncertainties introduced by experimental conditions increase uncertainties in prediction
- Experimental uncertainties:
  - geometry, initial and boundary conditions
  - material specifications - density, composition, grain structure

Inverse probability or inference



- Uncertainties introduced by experimental conditions increase uncertainties in inferred model parameters
- Experimental uncertainties:
  - geometry, initial and boundary conditions
  - material specifications - density, composition, grain structure

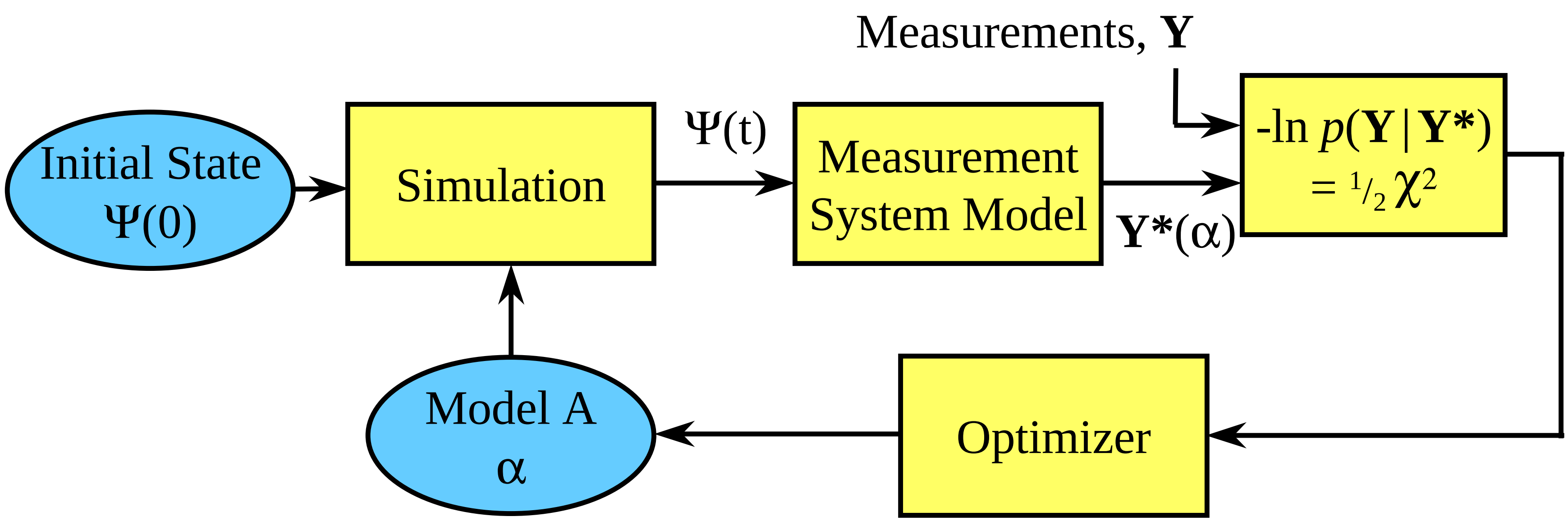
Comparison of simulation with experiment



Forward model

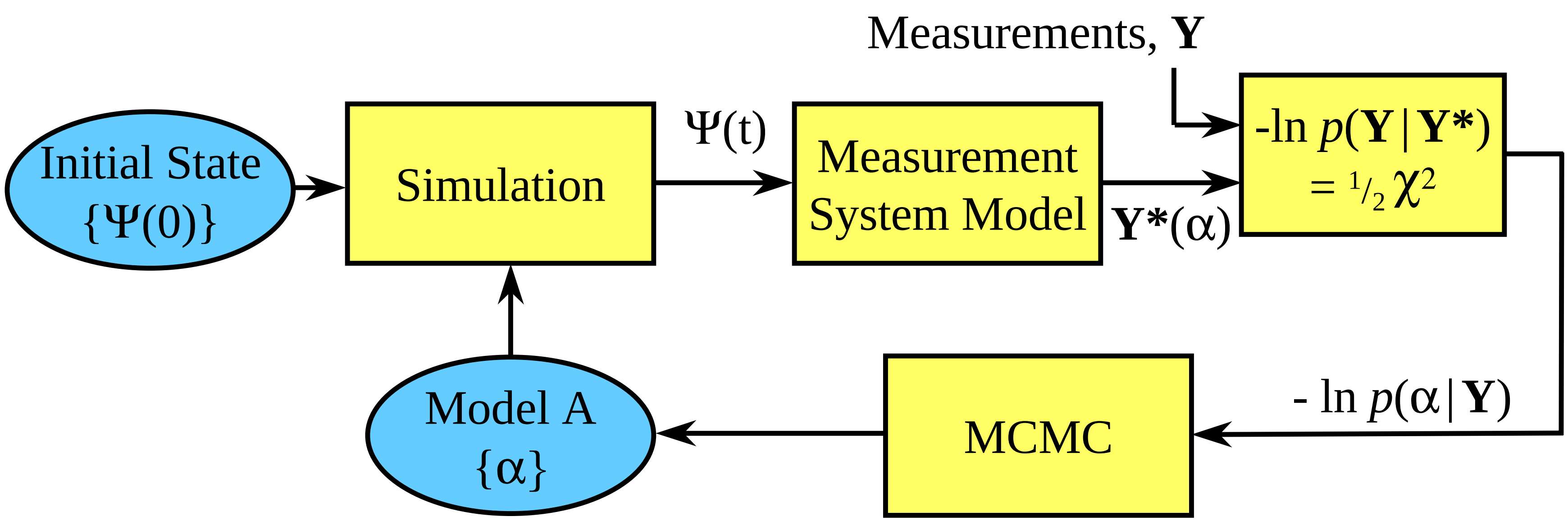
- Measurement system model transforms the simulated state of the physical system $\Psi(t)$ into measurements **Y*** that would be obtained in the experiment
- Mismatch with data summarized by minus-log-likelihood, $-\ln p(\mathbf{Y}\,|\,\mathbf{Y}^*) = {}^1\!/_2\,\chi^2$
- Inference - determine parameters from **Y**, $p(\alpha|\mathbf{Y})$ with Bayes law

Parameter estimation - maximum likelihood



- Optimizer adjusts parameters (vector α) to minimize -ln $p(\mathbf{Y}\,|\mathbf{Y}^*(\alpha))$
- Result is maximum likelihood estimate for α (also known as minimum-chi-squared solution)
- Optimization process is accelerated by using gradient-based algorithms along with adjoint differentiation to calculate gradients of forward model

Parameter uncertainties via MCMC



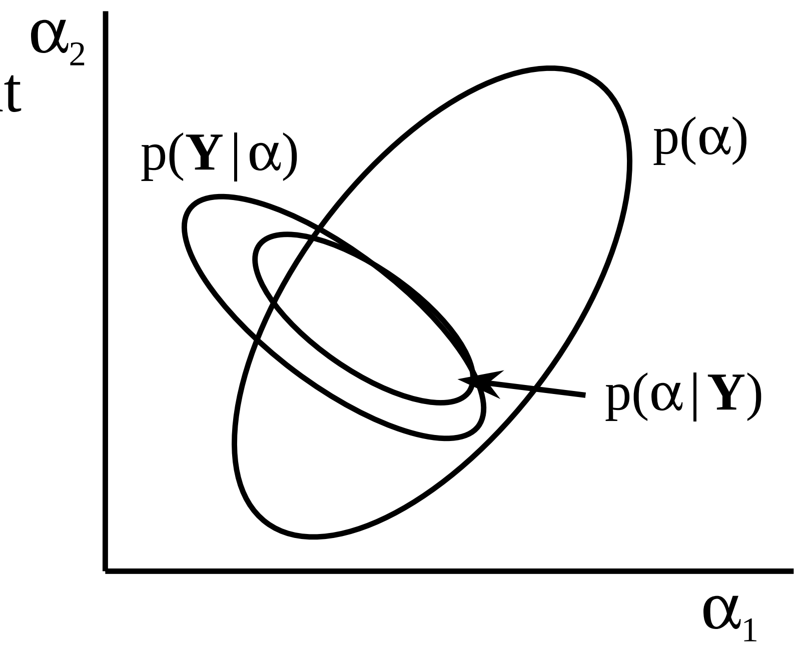
- Markov Chain Monte Carlo (MCMC) algorithm generates a random sequence of parameters that sample posterior probability of parameters for given data **Y**, $p(\alpha \mid \mathbf{Y})$, which yields plausible set of parameters $\{\alpha\}$.
- Must include uncertainty in initial state of system, $\{\Psi(0)\}$

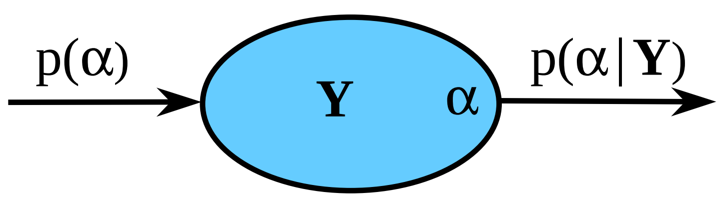
Analysis of many experiments involving several models

- Objective - combine results from many (all) experiments thereby reducing uncertainties in model parameters
  - include correlations among uncertainties, which are crucial but often neglected
  - require consistency of final models with all experiments
- Solution - link probabilistic analyses depicted by graphical representation
  - cumulative probabilistic analysis based on Bayes' law to optimally combine data
  - copes with complexity of analyzing large number of experiments
  - clearly displays logic and dependencies of analyses

Graphical probabilistic modeling

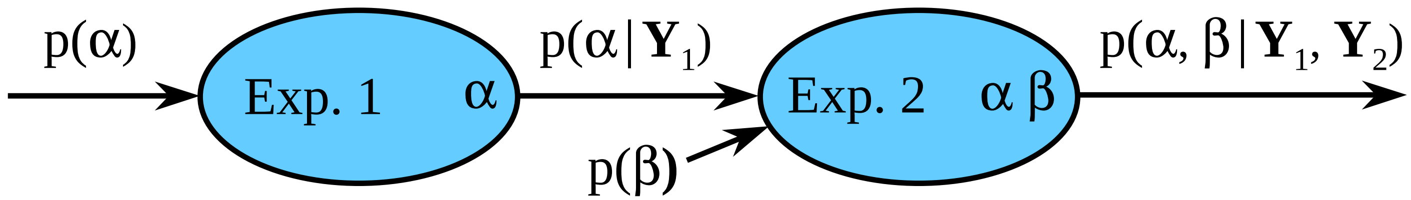
- Analysis of experimental data **Y** improves on prior knowledge about parameter vector α
- Bayes law:
  $$p(\alpha \mid \mathbf{Y}) \sim p(\mathbf{Y} \mid \alpha)\, p(\alpha)$$
  (posterior ~ likelihood x prior)
- Use bubble to represent effect of analysis based on data **Y**



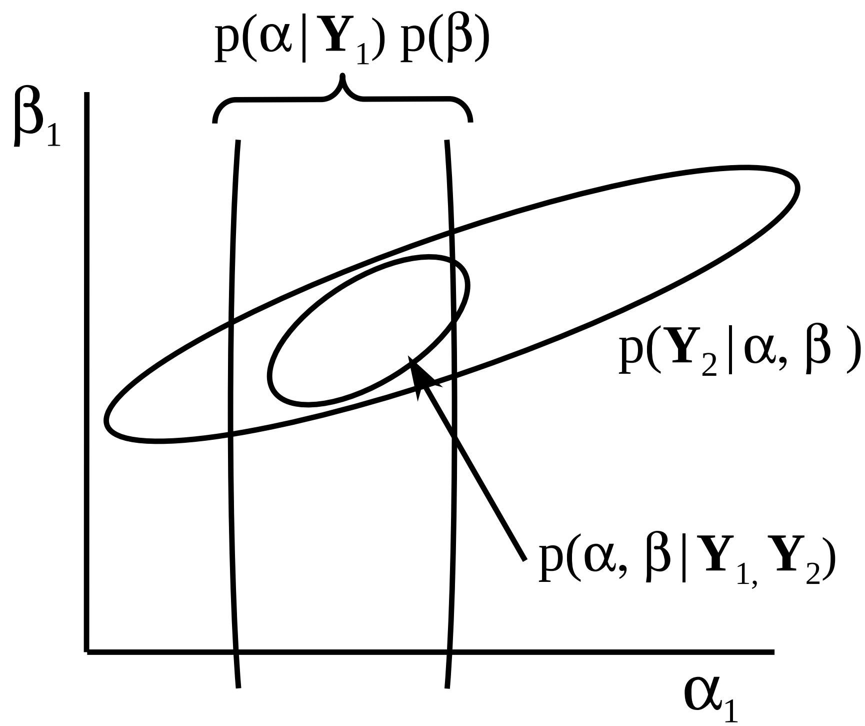


Graphical probabilistic modeling

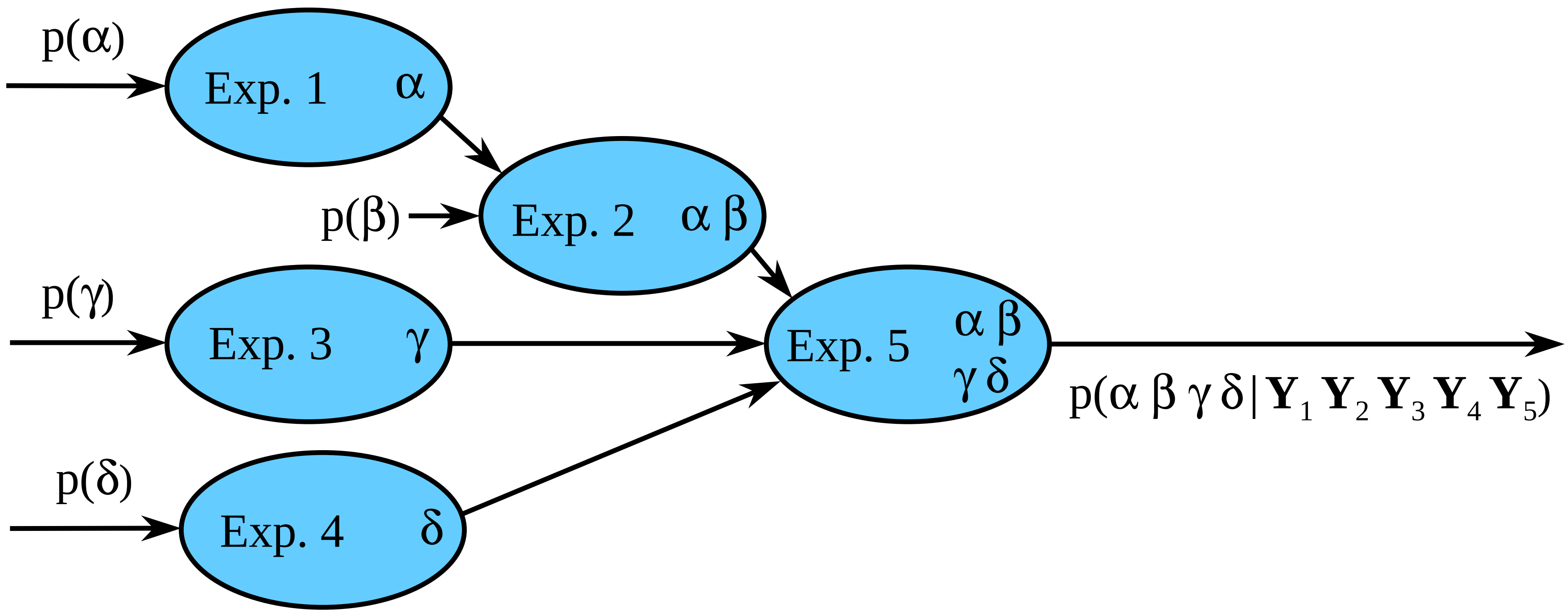
Propagate uncertainty through analyses of two experiments



- First experiment determines α, with uncertainties given by $p(\alpha|\mathbf{Y}_1)$
- Second experiment not only determines β but also refines knowledge of α
- Outcome is joint pdf in α and β, $p(\alpha, \beta|\mathbf{Y}_{1,}\mathbf{Y}_2)$ (NB: correlations)



Example of analysis of several experiments

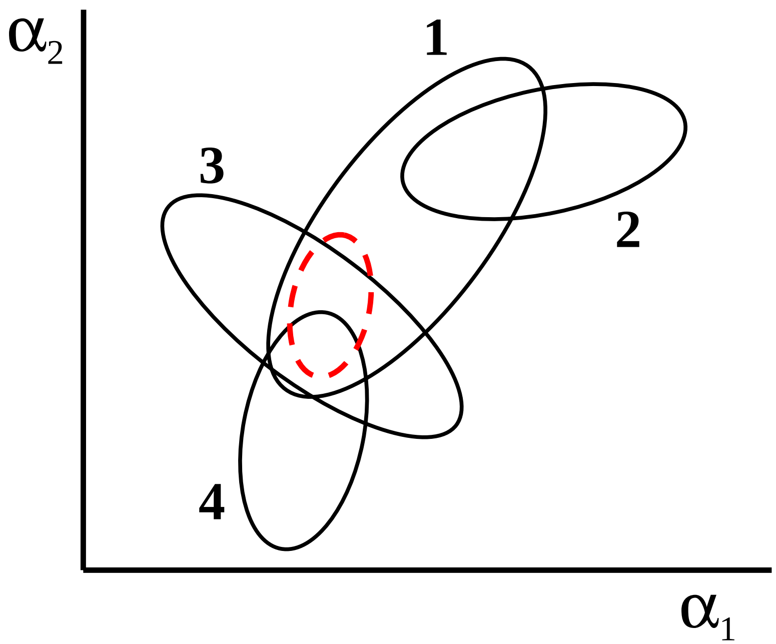


- Output of final analysis is full joint probability for all parameters based on all experiments
- Use of Gaussian pdfs simplifies process

Model checking

Check model consistent with all experimental data

- Important part of any analysis
- Check consistency of full posterior wrt. each of its contributions.
- Example shown at right:
  - likelihoods from Exps. 1 and 2 are mutually consistent
  - however, Exp. 2 is inconsistent with posterior (dashed) from all exps.
  - inconsistency must be resolved in terms of correction to model and/or interpretation of experiment



Graphical probabilistic modeling

- Diagrams useful for complete analysis of many experiments related to several models
  - displays logic
  - explicitly shows dependencies
  - organizational tool when many modelers and experimenters are involved
- Result is full joint probability for all parameters based on all previously analyzed experiments
  - uncertainties in all parameters, including their correlations, which are crucially important

Summary

- A methodology has been presented to combine experimental results from many experiments relevant to several basic physics models in the context of a simulation code
- Many challenges remain
  - ► systematic experimental uncertainties (effects common to many data)
  - ► identification and resolution of inconsistencies between experiments and simulation code
  - ► inclusion of other sources of uncertainty: material inhomogeneity, chaotic or turbulent behavior, numerical computation

Bibliography

- "A framework for assessing confidence in simulation codes," K. M. Hanson and F. M. Hemez, to appear in *Experimental Techniques* **25** (2001)
- "A framework for assessing uncertainties in simulation predictions," K. M. Hanson, *Physica D* **133**, pp. 179-188 (2000); an integrated approach to determining uncertainties in physics modules and their effect on predictions
- "Inversion based on complex simulations," K. M. Hanson, *Maximum Entropy and Bayesian Methods*, pp. 121-135 (Kluwer Academic, 1998); describes adjoint differentiation and its usefulness in simulation physics
- "Uncertainty assessment for reconstructions based on deformable models," K. M. Hanson et al., *Int. J. Imaging Syst. Technol.* **8**, pp. 506-512 (1997); use of MCMC to sample posterior

These and related papers available at http://home.lanl.gov/kmh/