Bayesian techniques for quantification of uncertainties

Ken Hanson

DX-3, Los Alamos National Laboratory

http://home.lanl.gov/kmh

Overview of presentation

- Introduction
  - uncertainty analysis
  - example - analysis of simple experiment
- General approach to analyzing single experiments
  - estimation of model parameters and uncertainty assessment
- Estimating uncertainties in simulation codes
- Graphical probabilistic modeling
  - analysis of numerous experiments in terms of many physical models
  - complete uncertainty analysis
- Summary

Uncertainty assessment and Stockpile Stewardship

- Presently we rely on experts to estimate the reliability of our weapons systems through the Enhanced Reliability Methodology
- Eventually we will need to rely more on hydrocodes to predict physical performance and safety of weapons
- A major challenge to NWT is to quantitatively assess uncertainties in hydrocode predictions
- Correctness of hydrocodes will be substantiated by comparison with legacy data and new non-nuclear experiments
- For uncertainty assessment, we need a carefully thought out methodology for combining old and new data and for designing new experiments to fill in missing information

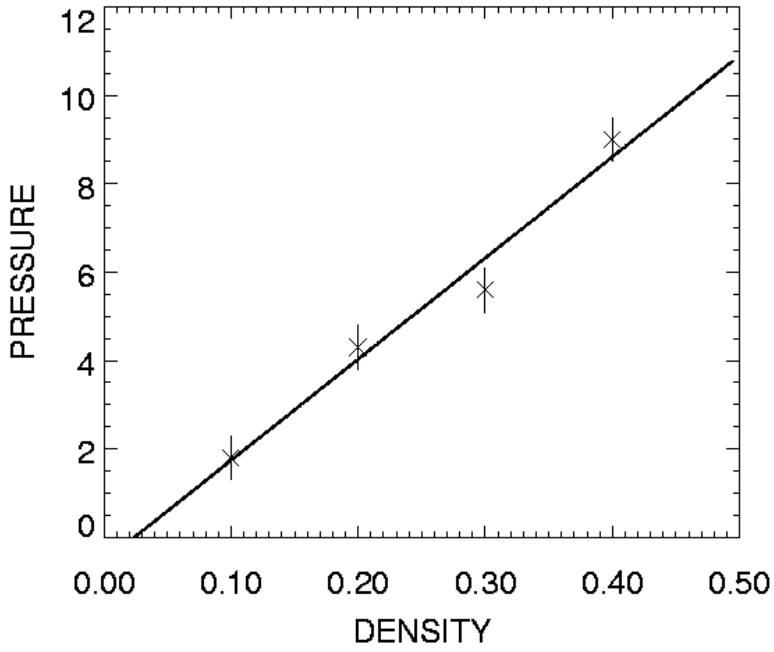
Uncertainty analysis

- Uncertainties in model parameters are characterized by probability density functions (pdf)
- Complete characterization of uncertainties in experiments
  - incorporate “systematic” uncertainties
  - include uncertainties in all experimental conditions
- Must include correlations among uncertainties
- Combine results from many (all) experiments
  - reduce uncertainties in model parameters
  - require consistency of models with all experiments

Isothermal dependence of gas pressure on density

Example of simple basic physics model

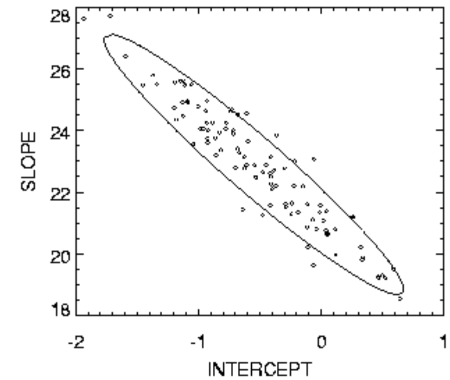
- Assume linear model to describe dependence (ideal gas)
- Determine two parameters, intercept and slope, by minimizing chi-squared based on four available measurements (standard "error" analysis)
- Use this linear model in simulation code where pressure of gas is needed and density is calculated

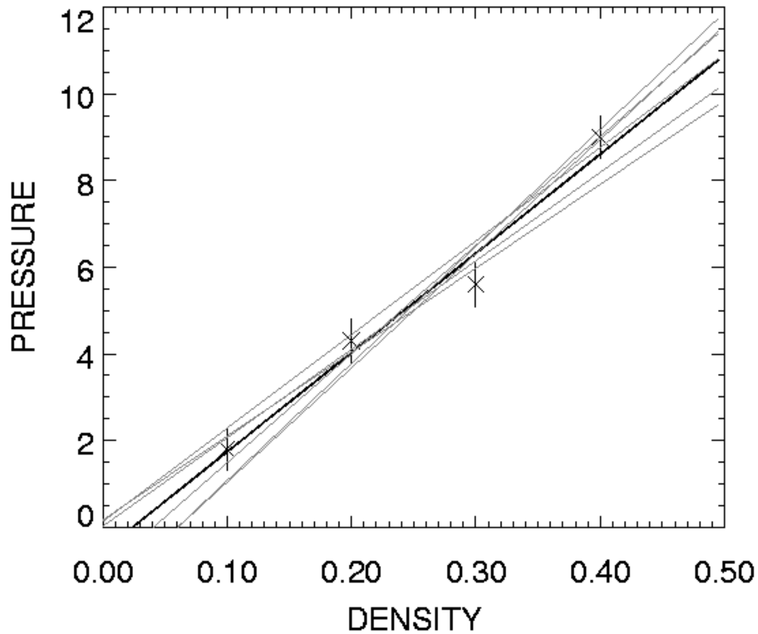


Isothermal dependence of gas pressure on density

Representation of uncertainties in inferred model

- Uncertainties in parameters, derived from uncertainties in measurements, given by Gaussian pdf in 2-D parameter space
  - correlations evidenced by tilt
  - points are random draws from pdf
- Should focus on implied uncertainties in physical phenomenon
  - light lines are plausible model realizations drawn from parameter pdf
  - characterize uncertainty in dependence of pressure vs. density

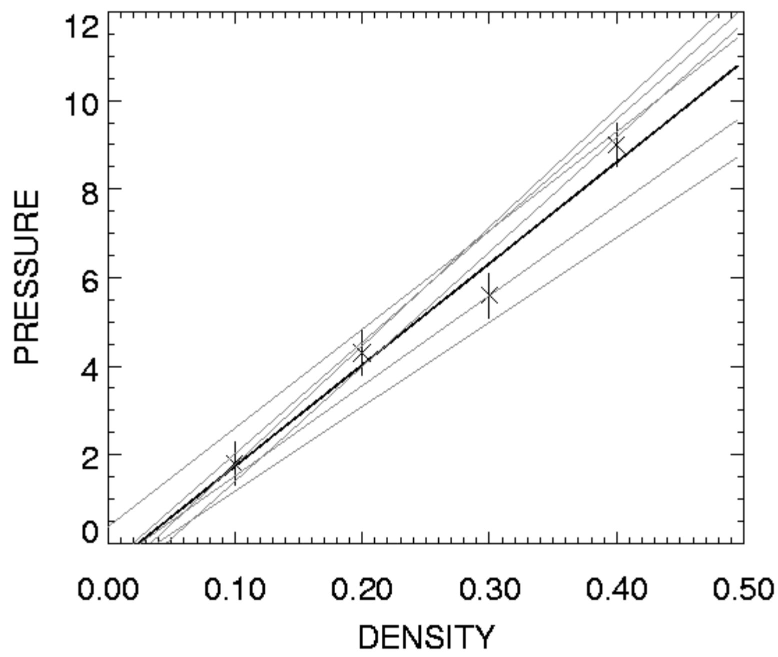




Isothermal dependence of gas pressure on density

Importance of correlations in uncertainties

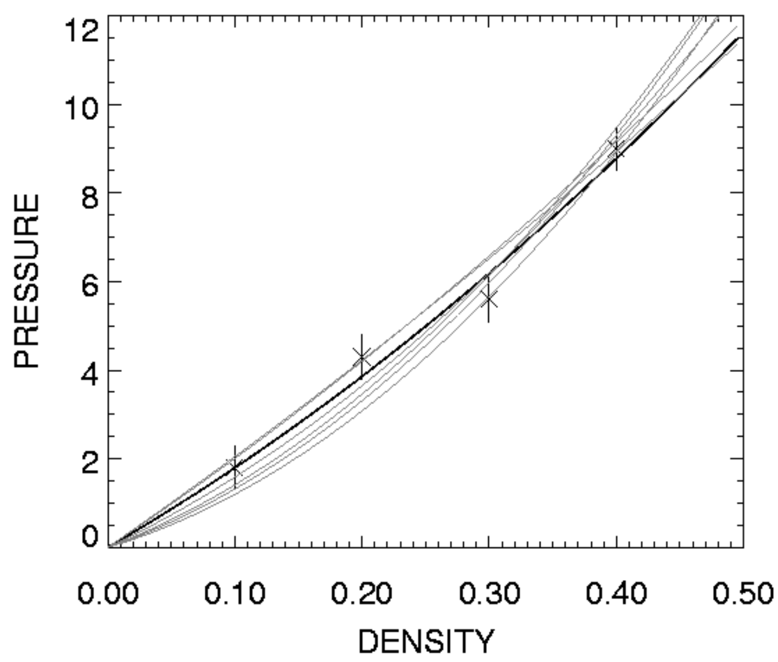
- Correlations in uncertainties are critically important
- Plot shows random samples from uncertainty in slope and intercept ignoring correlations
- Uncertainties in dependence of pressure vs. density **far exceed** uncertainties in measurements



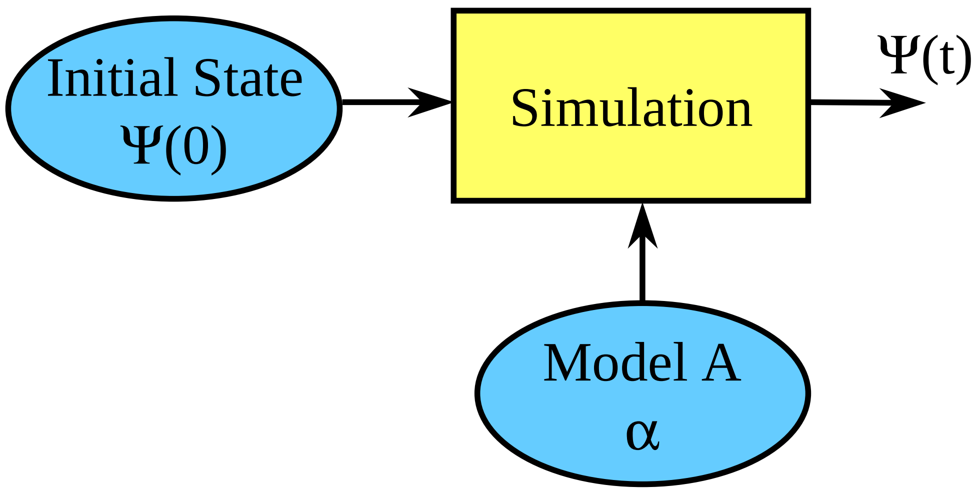
Isothermal dependence of gas pressure on density

Uncertainties in model affect overall uncertainty

- Quadratic might account for suspected departure from linearity
  - curve constrained to go through origin
- Comparison with previous linear model demonstrates increased uncertainties in model outside of density measurement range
- Conclusion: basic physics experiments should cover full operating range of physical variables used by simulation code; extrapolation increases uncertainty

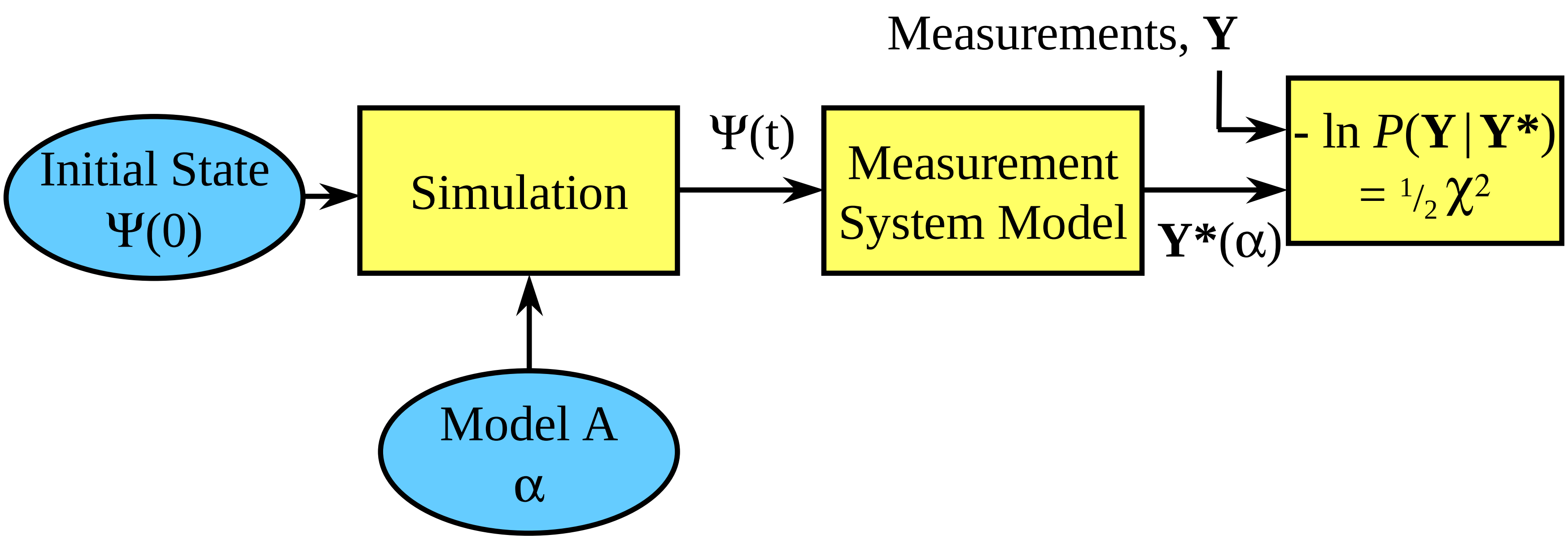


Simulation code



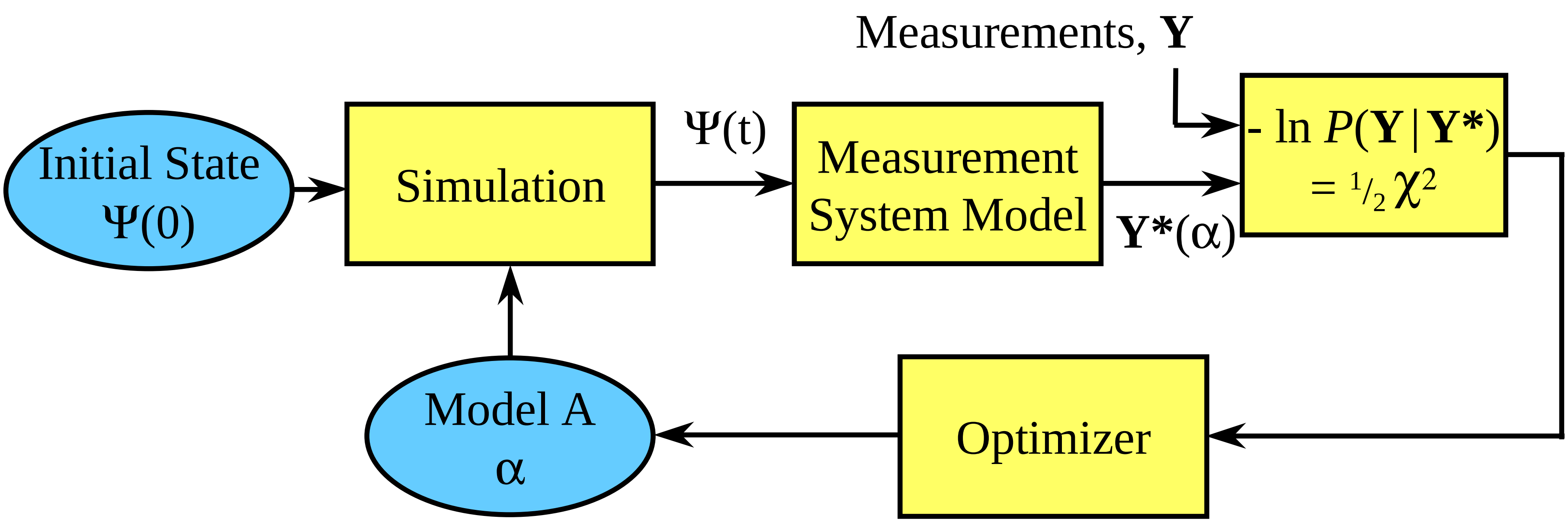
- Simulation code predicts state of time-evolving system
  - $\Psi(t)$ = time-dependent state of system
  - $\Psi(0)$ = initial state of system
- Properties of one system component described by physics model A with parameter vector α

Comparison of simulation with experiment



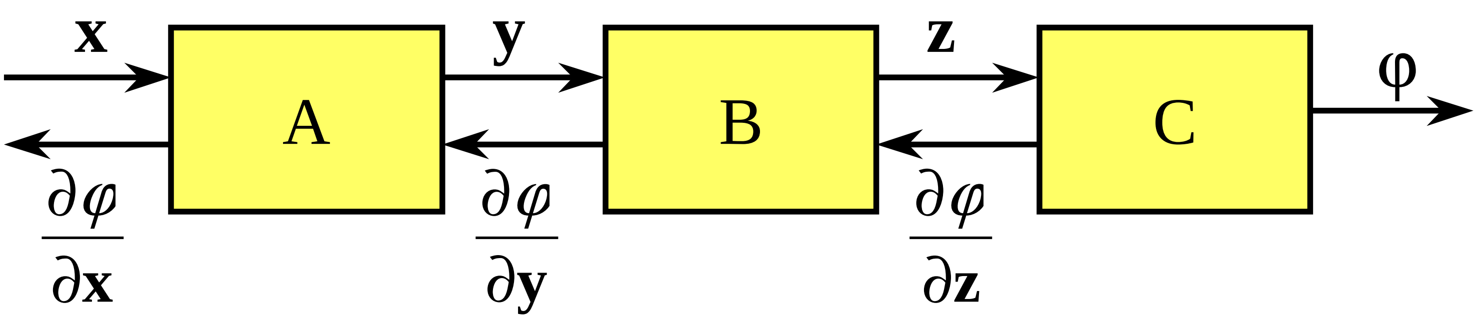
- Measurement system model transforms the simulated state of the physical system $\Psi(t)$ into measurements $\mathbf{Y^*}$ that would be obtained in the experiment
- Mismatch to data summarized by minus-log-likelihood, $-\ln P(\mathbf{Y} \,|\, \mathbf{Y^*}) = {}^1/_2\, \chi^2$

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 together with adjoint differentiation to calculate gradients of forward model

Adjoint Differentiation of Forward Calculation



- Data-flow diagram shows sequence of transformations A, B, C that convert data structure **x** to **y** to **z** and then scalar φ.
- Derivatives of φ with respect to **x** are efficiently calculated in the reverse (adjoint) direction.
- CPU time to compute **all** derivatives is comparable to forward calculation
- One may need to keep intermediate data structures to evaluate derivatives
- Code based: logic of adjoint code derivable from forward code

Analysis of single experiment

- Likelihood
  - p(**Y**|**Y***) = probability of measurements **Y** given the values **Y*** predicted by experiment simulation. (NB: **Y*** depends on α)
- The pdf describing uncertainties in model parameter vector α, called posterior:
  - $p(\alpha\,|\,\mathbf{Y}) \propto p(\mathbf{Y}\,|\,\mathbf{Y}^*)\, p(\alpha)$ (Bayes law)
  - $p(\alpha)$ is prior; summarizes previous knowledge of α
  - "best" parameters estimated by maximizing $p(\alpha|\mathbf{Y})$ (called Maximum A Posteriori solution)
  - uncertainties in α are fully characterized by $p(\alpha|\mathbf{Y})$

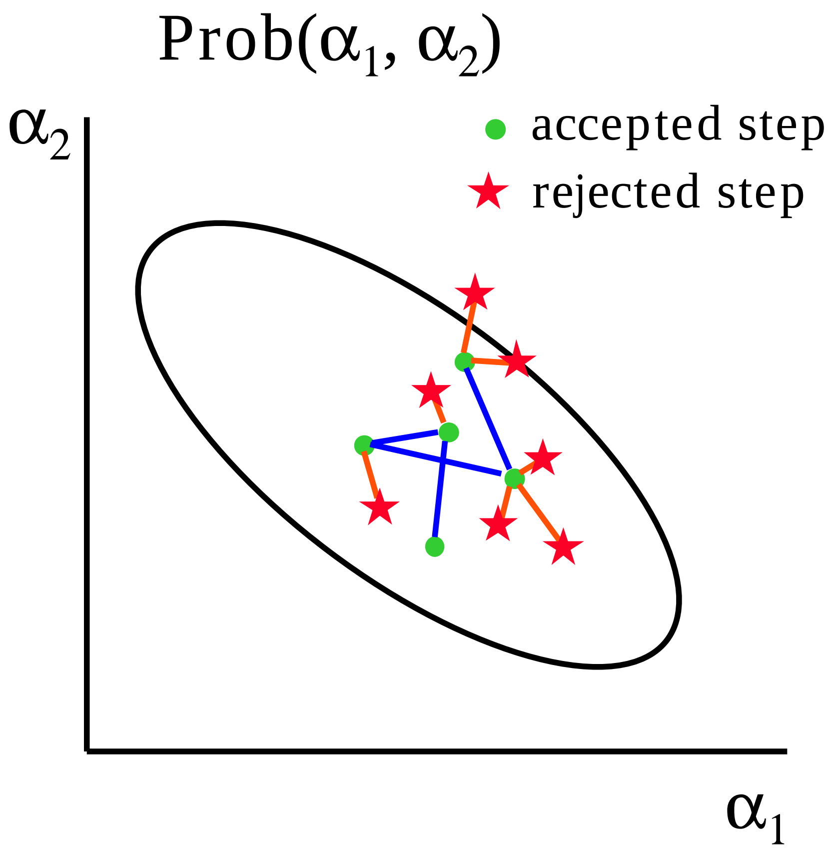
Parameter uncertainties via MCMC

- Posterior $p(\alpha|\mathbf{Y})$ provides full uncertainty distribution for α
- Markov Chain Monte Carlo (MCMC) algorithm generates a random sequence of parameters that sample $p(\alpha|\mathbf{Y})$
  - results in plausible set of parameters $\{\alpha\}$
  - representative of uncertainties
  - second moments of parameters can be used to estimate covariance matrix **C**
- MCMC advantages
  - can be applied to any pdf, not just Gaussians
  - automatic marginalization over nuisance variables
- MCMC disadvantage
  - potentially calculationally demanding

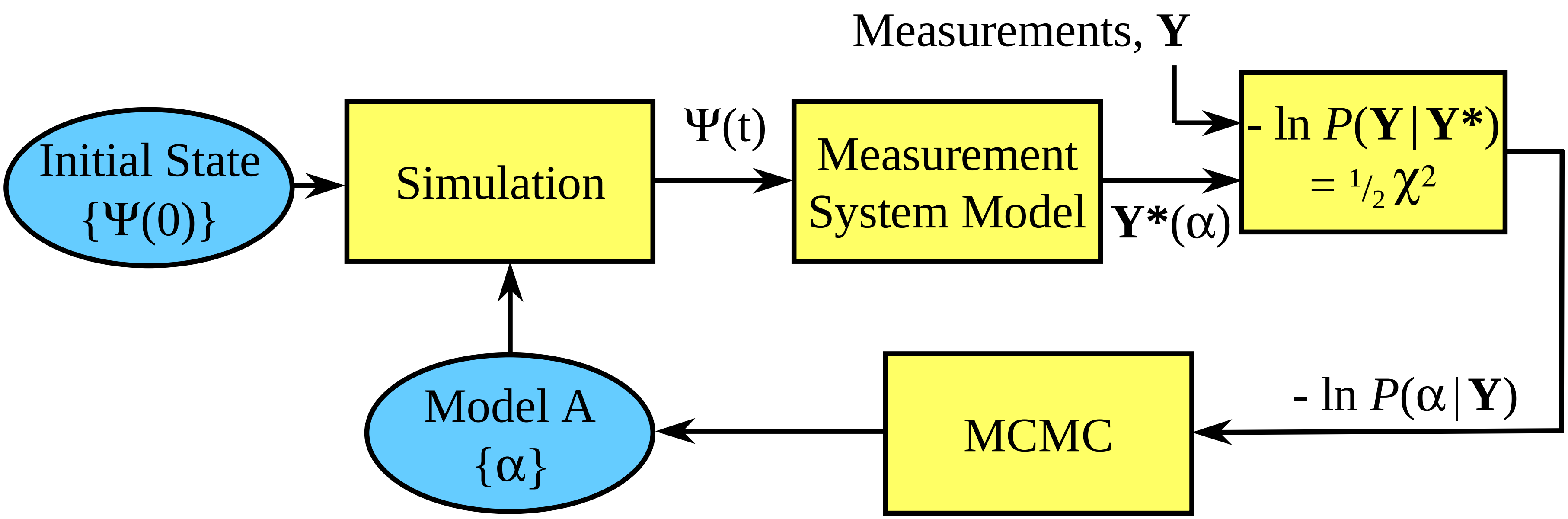
Markov Chain Monte Carlo

Generates sequence of random samples from an arbitrary probability density function

- Metropolis algorithm:
  - draw trial step from symmetric pdf, i.e., $T(\Delta\alpha) = T(-\Delta\alpha)$
  - accept or reject trial step
  - simple and generally applicable
  - relies only on calculation of target pdf for any α
  - works well for many parameters



Parameter uncertainties via MCMC



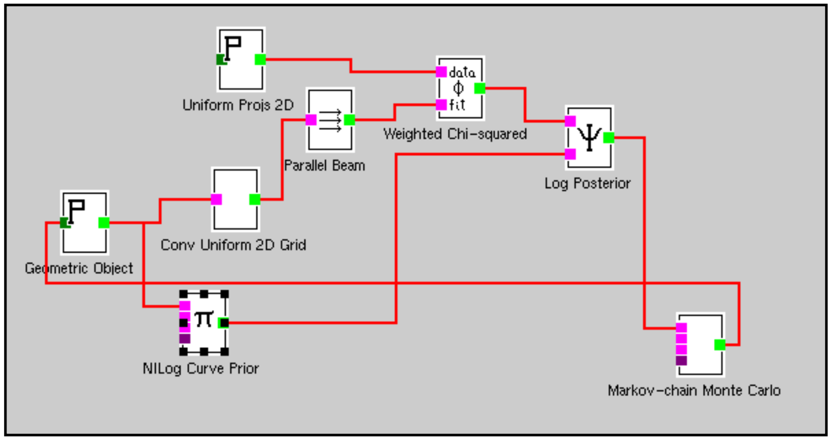
- Markov Chain Monte Carlo (MCMC) algorithm generates a random sequence of parameters that sample posterior probability of parameters for given data **Y,** $P(\alpha|\mathbf{Y})$, which yields plausible set of parameters $\{\alpha\}$.
- Must include uncertainty in initial state of system, $\{\Psi(0)\}$

Uncertainty analysis with Bayes Inference Engine

Represents general paradigm for analyzing an experiment

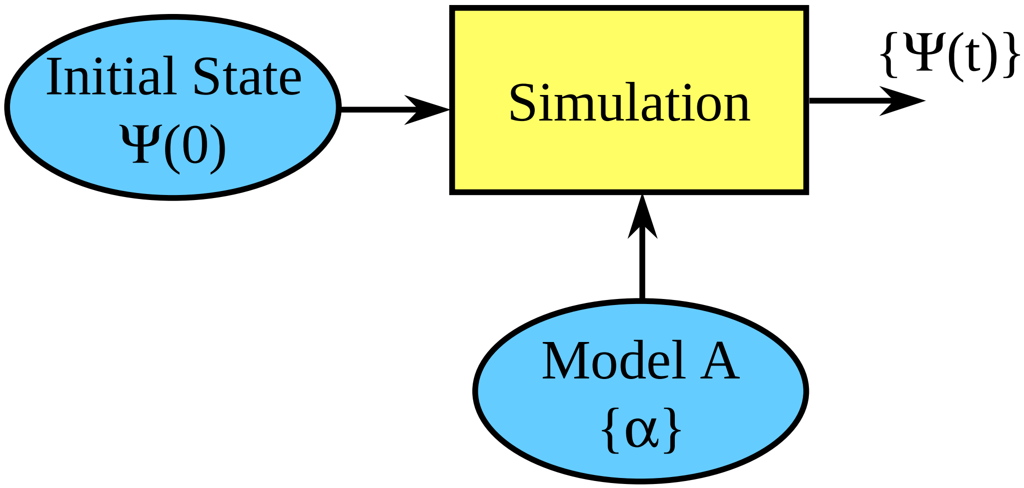
- Problem of reconstruction from just two radiographs solved with Bayes Inference Engine (BIE) using deformable boundary
- Markov Chain Monte Carlo generates set of plausible solutions, whose fluctuations characterize uncertainty in boundary localization, by drawing random samples from posterior probability distribution

Data flow diagram in BIE



Reconstruction with several plausible boundaries

Simulation of plausible predictions - characterizes uncertainty in prediction



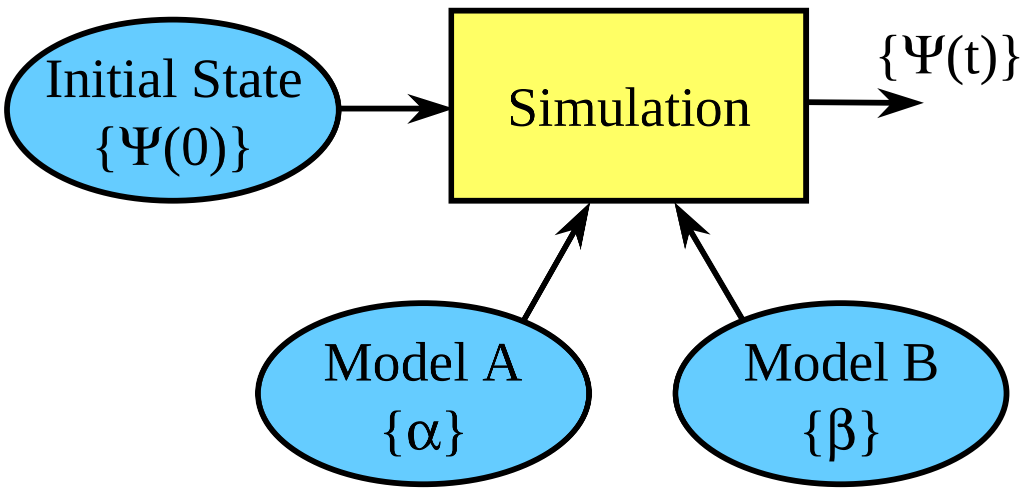
- Generates plausible predictions for known uncertainties in parameters
  - $\{\Psi(t)\}$ = plausible sets of dynamic state of system
  - $\{\alpha\}$ = plausible sets of parameter vector α
- Use Monte Carlo method - run simulation code for each random draw from posterior for α, $P(\alpha|.)$, to obtain set of predictions $\{\Psi(t)\}$

Uncertainty in predictions

- Estimate by propagating through simulation code a set of 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 in simulation may be included later
  - variability in densities
  - chaotic behavior

UNCLASSIFIED

Plausible outcomes for many models

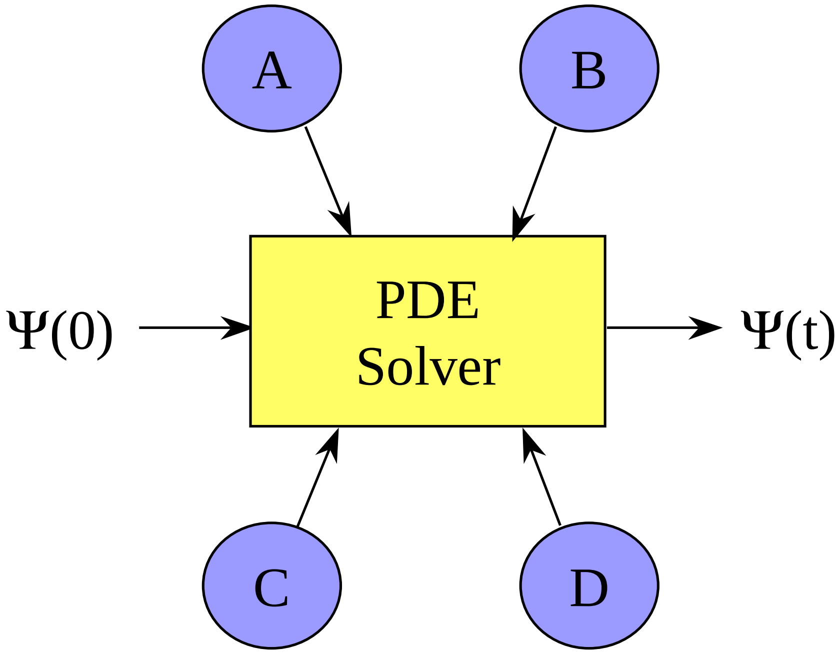


- Integrated simulation code predicts plausible results for known uncertainties in initial conditions and material models
  - {Ψ(t)} = plausible sets of dynamic state of system
  - {Ψ(0)} = plausible sets of initial state of system
  - {α} = plausible sets of parameter vector α for material A
  - {β} = plausible sets of parameter vector β for material B

Simulation code

Predicts state of time-evolving physical system

- Future state of system Ψ(t) predicted from initial state Ψ(0)
- Code consists of two parts:
  - PDE solver computes basic dynamical equations
  - models for material behavior
- To determine accuracy of Ψ(t)
  - determine accuracy of PDE solver (verification)
  - determine effects of uncertainties in material models on Ψ(t) (validation)



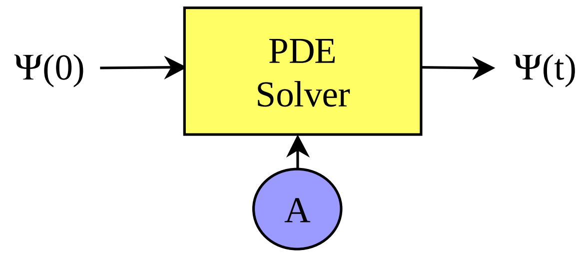
Uncertainty in simulation predictions

- 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 in simulation may be included later
  - variability in densities
  - chaotic behavior

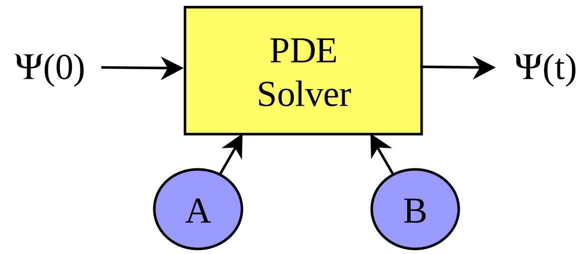
Validation Experiments

Full validation requires hierarchy of experiments

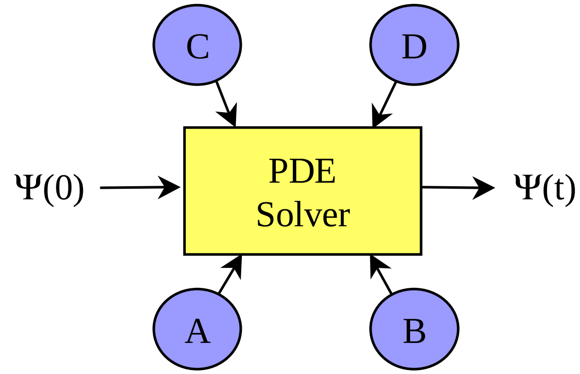
- **Basic** experiments determine elemental physics models



- **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

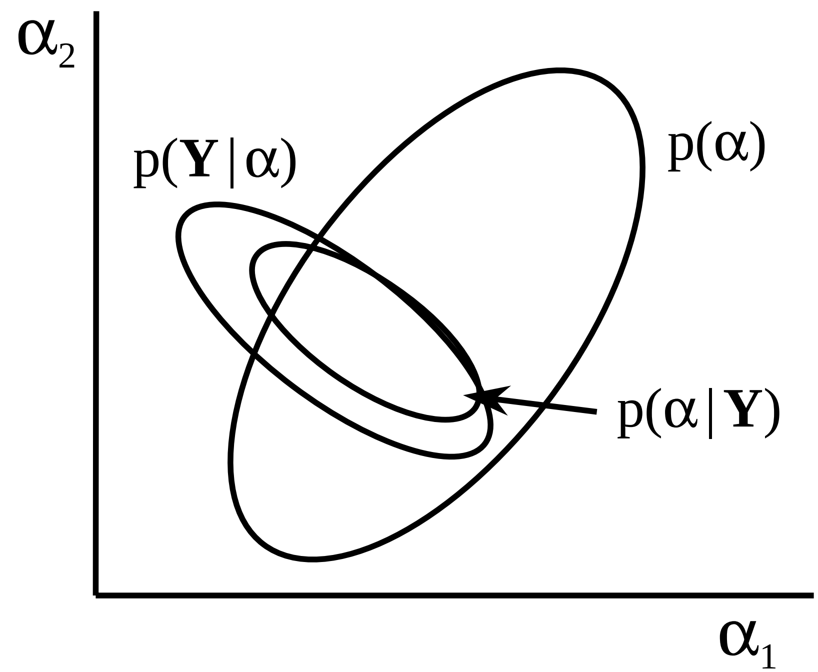


Analysis of many experiments involving several models

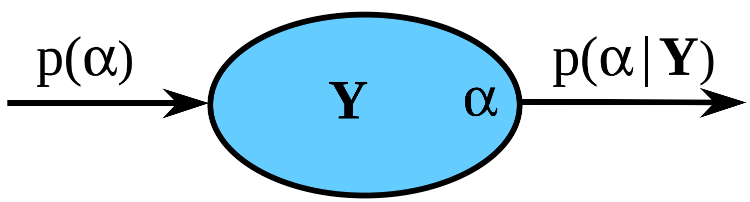
- Objective - combine results from many (all) experiments thereby reducing uncertainties in model parameters
  - include correlations among uncertainties, which are crucial but often neglected
  - require consistency of final models with all experiments
- Solution - link probabilistic analyses depicted 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

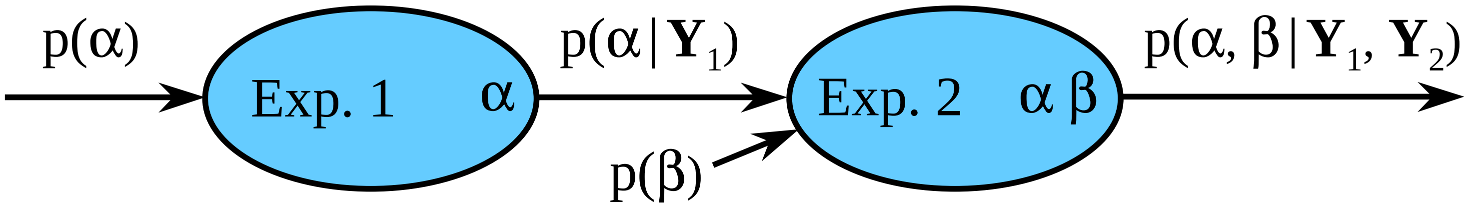
- Analysis of experimental data **Y** improves on prior knowledge about parameter vector α
- Bayes law:
  $p(\alpha|\mathbf{Y}) \sim p(\mathbf{Y}|\alpha)\, p(\alpha)$
  (posterior ~ likelihood x prior)



- Use bubble to represent effect of analysis based on data **Y**
- In terms of logs: $-\ln p(\alpha|\mathbf{Y}) = -\ln p(\mathbf{Y}|\alpha) - \ln p(\alpha) + \text{constant}$



Graphical probabilistic modeling



Output of second bubble:

$$p(\alpha, \beta \,|\, \mathbf{Y}_1, \mathbf{Y}_2) \sim p(\mathbf{Y}_1, \mathbf{Y}_2 \,|\, \alpha, \beta)\; p(\alpha, \beta) \qquad \text{(Bayes law)}$$

$$\sim p(\mathbf{Y}_2 \,|\, \alpha, \beta)\; p(\beta)\; p(\alpha \,|\, \mathbf{Y}_1)$$

(likelihood 2 x prior(β) x posterior 1)

$$\sim p(\mathbf{Y}_2 \,|\, \alpha, \beta)\; p(\beta)\; p(\mathbf{Y}_1 \,|\, \alpha)\; p(\alpha)$$

(likelihood 2 x prior(β) x likelihood 1 x prior(α))

Summary: Action of bubble is to multiply input pdfs by likelihood from experiment to get output joint pdf

Use of logarithms of probabilities

- In terms of log-probability, Bayes law becomes:

  $$- \ln p(\alpha|\mathbf{Y}) = - \ln p(\mathbf{Y}|\alpha) - \ln p(\alpha) + \text{constant}$$

- Parameters are estimated by minimizing $- \ln p(\alpha|\mathbf{Y})$
- Gaussian approximation of probability:

  $$- \ln p(\alpha) = \phi = \phi_0 + (\alpha - \alpha_0)^T \mathbf{K} (\alpha - \alpha_0) \,,$$

  where $\mathbf{K}$ is the curvature or second derivative matrix of ϕ (aka Hessian) and α_0 is the position of the minimum in ϕ
- Covariance matrix, $\mathbf{C} = \langle(\alpha - \alpha_0)(\alpha - \alpha_0)^T\rangle$,

  is inverse of $\mathbf{K}$: $\mathbf{C} = \mathbf{K}^{-1}$
- Likelihood for Gaussian measurement uncertainties is

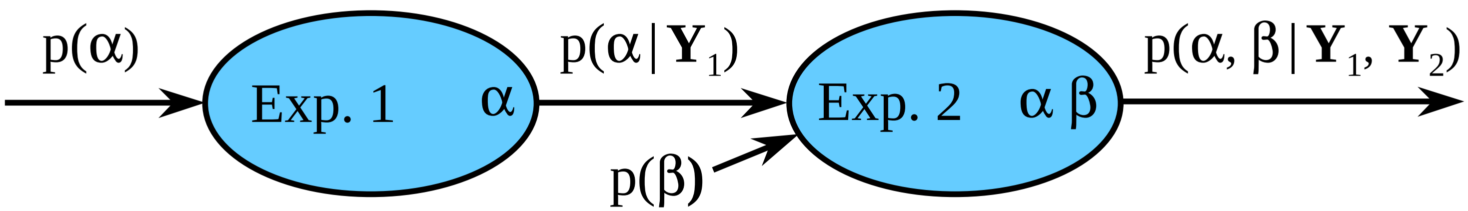
  $$-\ln P(\mathbf{Y}|\mathbf{Y^*}) = {}^1/_2\, \chi^2 = {}^1/_2 \sum \{(y_i - y^*_i)/\sigma_i\}^2$$

Gaussian probabilities
Lead to simplified combination of pdfs

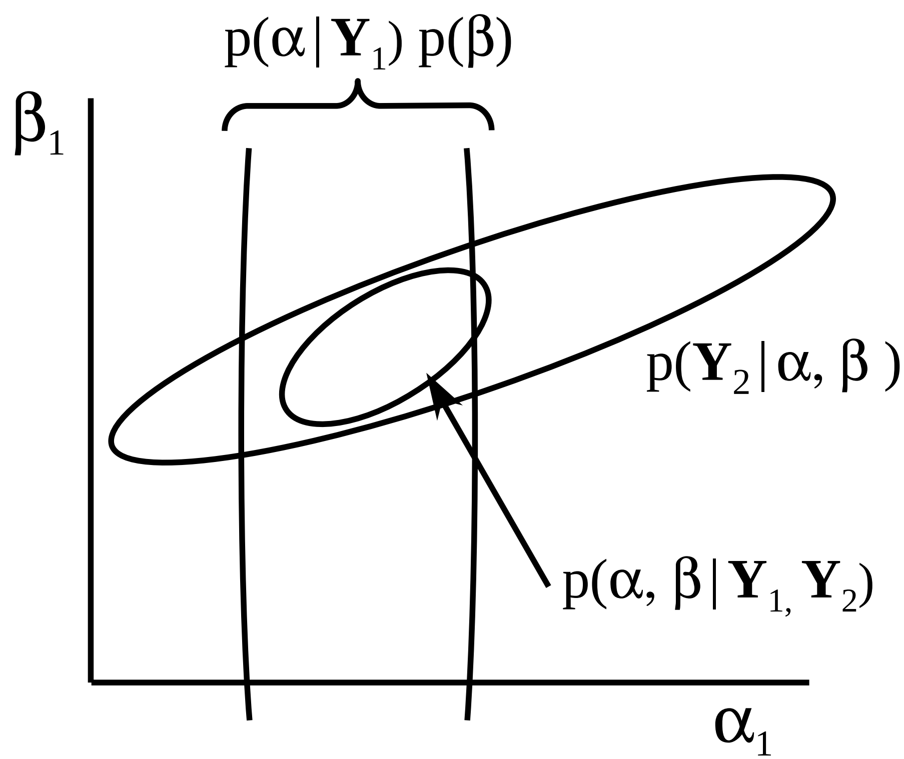
- In terms of log-probability, Bayes law, gives posterior as:
  $$- \ln p(\alpha|\mathbf{Y}) = - \ln p(\mathbf{Y}|\alpha) - \ln p(\alpha) + \text{constant}$$
- Bayes law for Gaussians,
  $$- \ln p(\alpha|\mathbf{Y}) = \phi = \phi_0 + (\alpha - \alpha_0)^T \mathbf{K_0} (\alpha - \alpha_0) =$$
  $$(\alpha - \alpha_\mathbf{L})^T \mathbf{K_L} (\alpha - \alpha_\mathbf{L}) + (\alpha - \alpha_\mathbf{P})^T \mathbf{K_P} (\alpha - \alpha_\mathbf{P}) + \text{const.},$$
  where subscripts L & P correspond to likelihood & prior
- Curvature matrix of posterior is: $\mathbf{K_0} = \mathbf{K_L} + \mathbf{K_P}$
- Covariance matrix of posterior is: $\mathbf{C_0} = \mathbf{K_0}^{-1} = [\mathbf{K_L} + \mathbf{K_P}]^{-1}$
- Estimated parameters are: $\alpha_0 = \mathbf{K_0}^{-1} [\alpha_\mathbf{L} \mathbf{K_L} + \alpha_\mathbf{P} \mathbf{K_P}]$

Graphical probabilistic modeling

Propagate uncertainty through a sequence of analyses



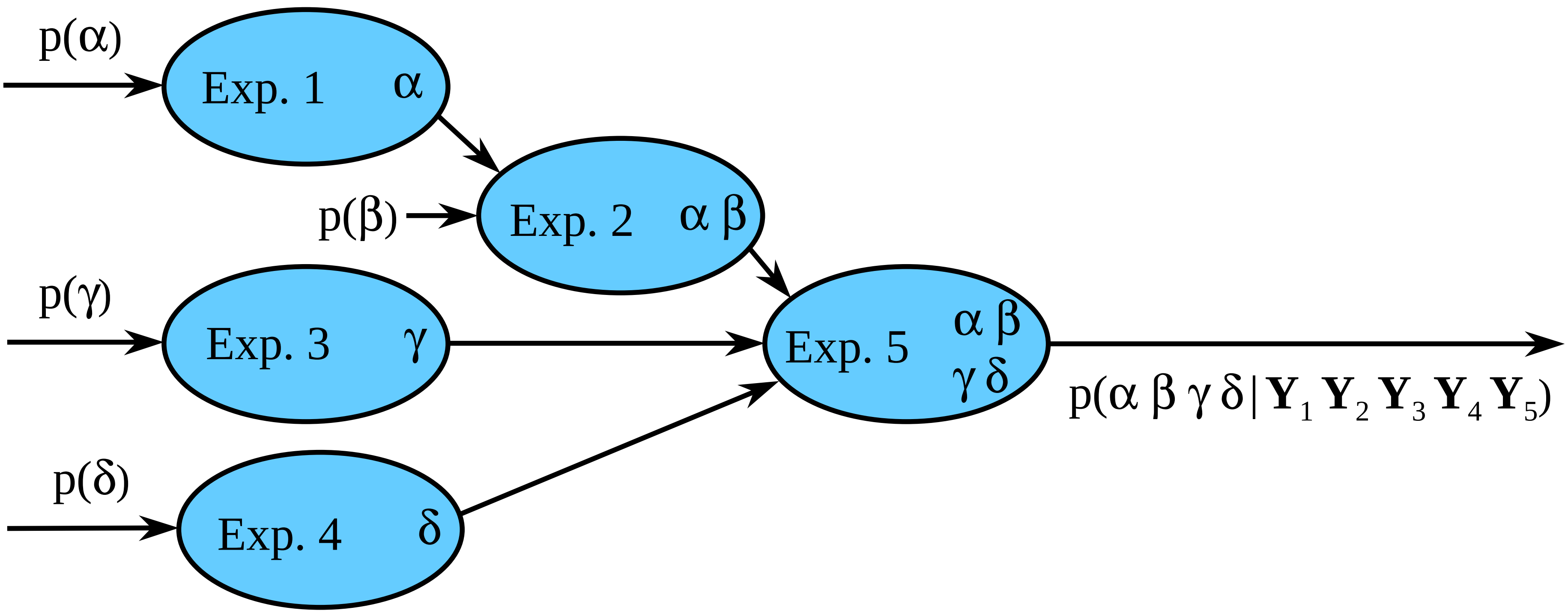
- First experiment determines α, with uncertainties given by $p(\alpha|\mathbf{Y}_1)$
- Second experiment not only determines β but also refines knowledge of α
- Outcome is joint pdf in α and β, $p(\alpha, \beta|\mathbf{Y}_{1,}\mathbf{Y}_2)$ (NB: correlations)



Graphical probabilistic modeling

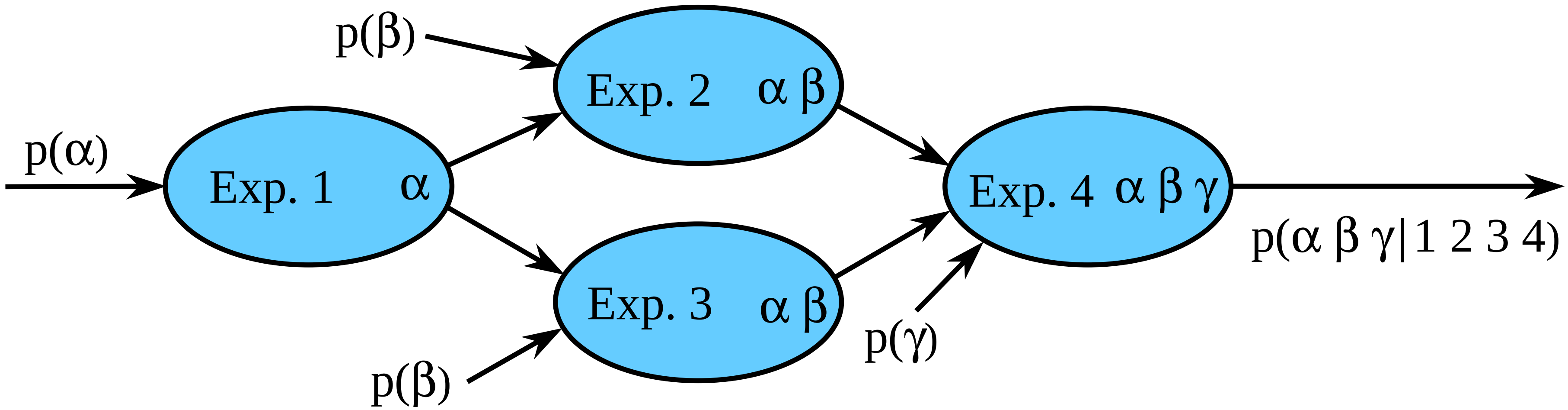
- Diagrams useful for complete analysis of many experiments related to several models
  - displays logic
  - explicitly shows dependencies
  - organizational tool when many modelers and experimenters are involved
- Result is full joint probability for all parameters based on all previously analyzed experiments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

Example of analysis of several experiments



Output of final analysis is full joint probability for all parameters based on all experiments

Need to avoid double counting of data



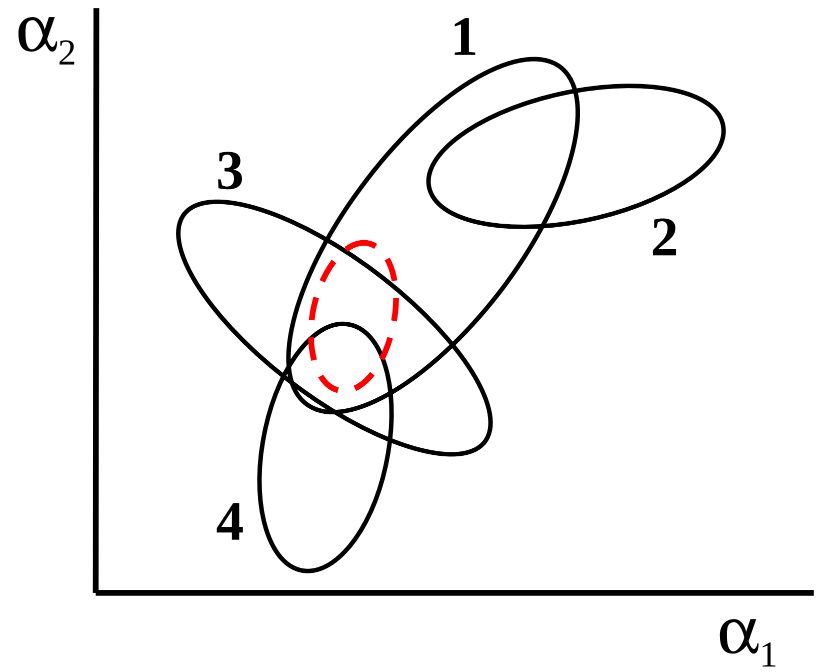
Outputs of analyses of both Exps. 2 and 3 make use of output of Expt. 1 and prior on β. This repetition must be avoided in overall posterior calculation through dependency analysis:

$$- \ln p(\alpha\ \beta\ \gamma | 1\ 2\ 3\ 4) = - \ln p(1|\alpha) - \ln p(\alpha) - \ln p(2|\alpha\ \beta) - \ln p(\beta)$$
$$- \ln p(3|\alpha\ \beta) - \ln p(4|\alpha\ \beta\ \gamma) - \ln p(\gamma) + \text{constant}$$

Model checking

Check that model is 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 consistent with each other
  - 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



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)
  - detection and resolution of inconsistencies between experiments and simulation code
  - inclusion of other sources of uncertainty: material inhomogeneity, chaotic or turbulent behavior

Proposal

- Propose building application to implement this approach
  - database of experiments showing links between analyses
  - logically consistent inferences about models based on all information
  - natural way to understand limits to parameter adjustment to match data from fully integrated experiments