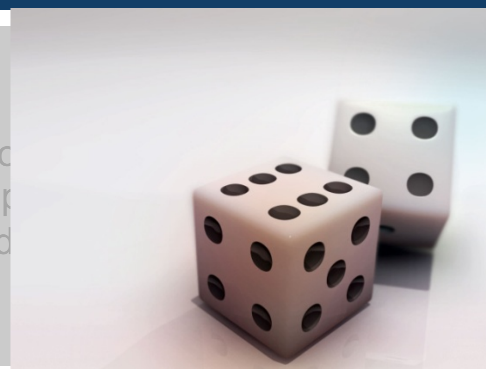
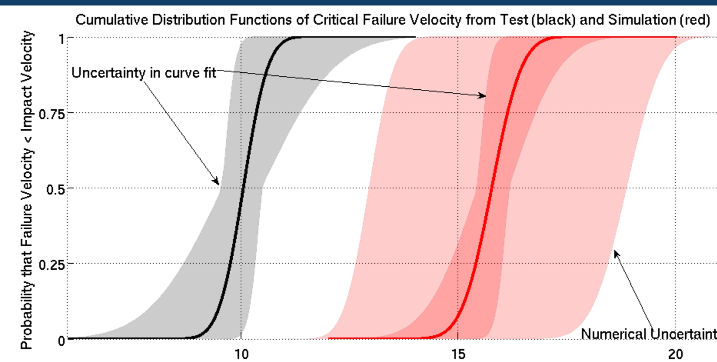
SAND2015-3623C

Exceptional service in the national interest

Sandia National Laboratories





A (more) Holistic Approach to Uncertainty Estimation

Bill Rider, Sandia National Laboratories, Albuquerque, Center for Computing Research

U.S. DEPARTMENT OF ENERGY
NNSA National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2011-XXXXP

“Most daily activity in science can only be described as tedious and boring, not to mention expensive and frustrating.”

Stephen J. Gould, Science, Jan 14, 2000.

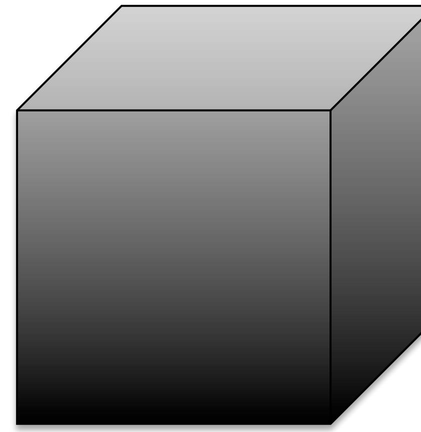
Major Issues and Points Discussed

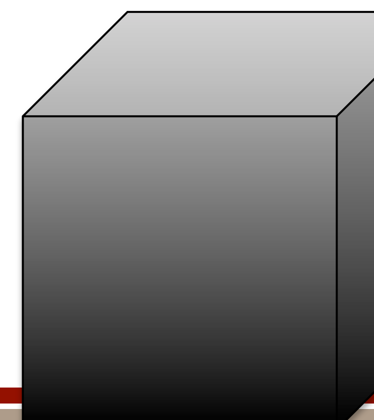
- Categories of uncertainty
  - Epistemic
  - Aleatory
  - Numerical
  - User
- Black boxes versus white boxes
- Where effort is needed, where effort is valuable
- Combining Uncertainties

Uncertainty estimation will always be incomplete, but should it be as incomplete as it is now?

Epistemic uncertainty is the current focus largely because it is straight forward to compute using existing tool.

- Epistemic Uncertainty – associated with lack of knowledge, hence it is reducible.
- In some cases it is simply thought of as uncertainty
- In many cases it is measured via looking at modeling variations
- Tools exist to examine modeling to the extent that model parameters can be modified simply.
- Sometimes the model parameters are chosen in regimes that the model should never occupy.





Aleatory is completely under-estimated to ignored because tools are not available

- Definition – this is the random aspect of uncertainty.
  - ala dice rolling
  - This is intrinsic and cannot be reduced
  - Referred to variability by some

- The line between epistemic and aleatory is fuzzy
- For example some of the randomness is due to inability to measure the conditions being simulated.
  - If conditions could be measured accurately enough the computation would be completely deterministic (?)

Epistemic and aleatory have fuzzy distinctions (Hills)

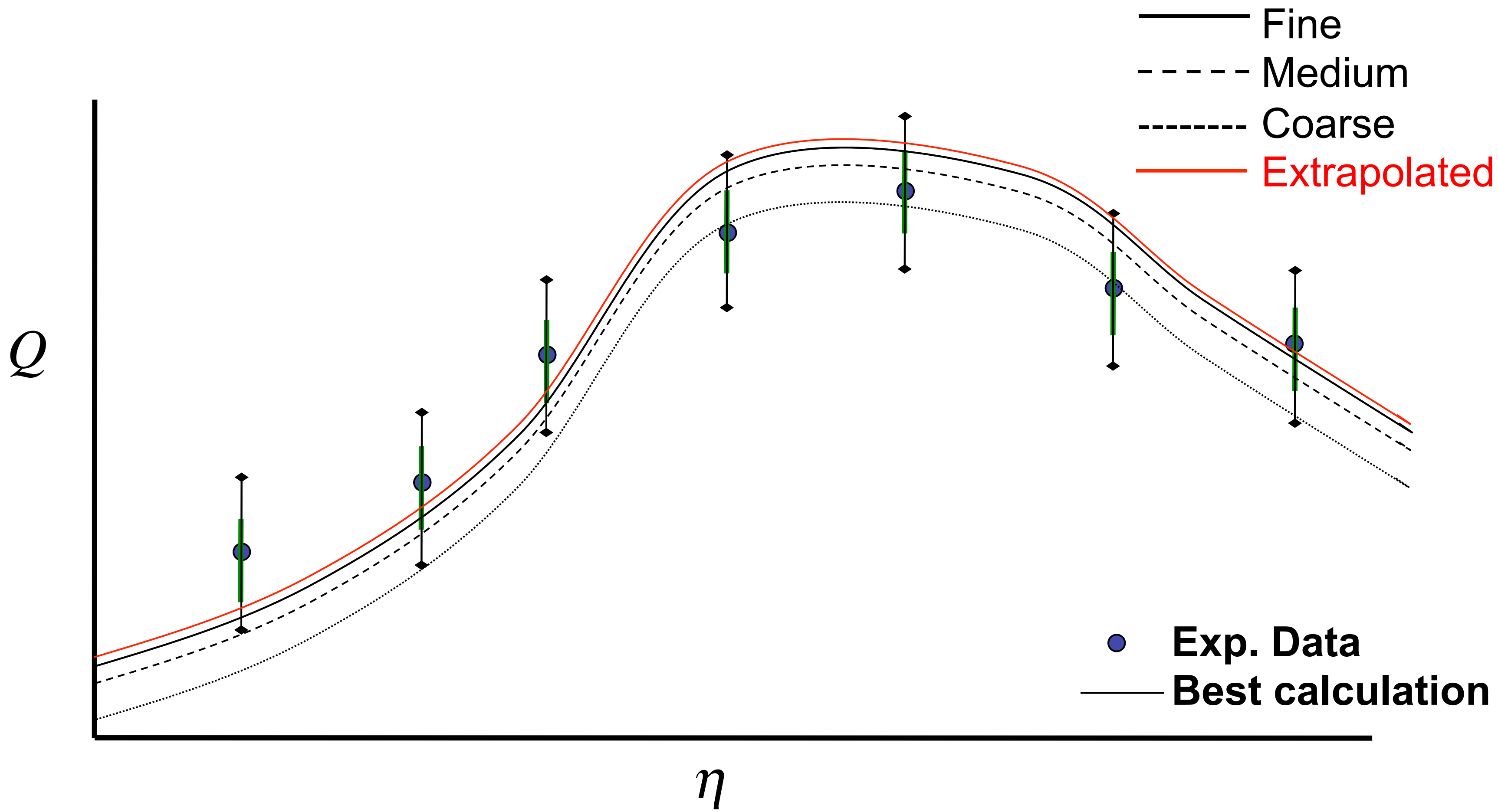
Some other conflicting opinions (all have merit):

- There is no such thing as aleatory uncertainty. All uncertainty is due to lack of knowledge
- There is no need or benefit to separating aleatory and epistemic uncertainty
- It is possible to use subjective probability to characterize expert opinion
- The use of subjective probability to characterize expert opinion is likely to produce garbage in / garbage out

Our approach:

- Based on well established techniques for high consequence applications developed by the EPA/NRC for quantitative/probabilistic risk assessment
- The approach separates epistemic and aleatory uncertainty
- We should use care and understand the limitations in the elicitation of expert opinion
- We should understand the assumptions and limitations of our analysis and be transparent as to the limitations and assumptions made in applying these techniques

The experimental "error" has two components (observation & variability).

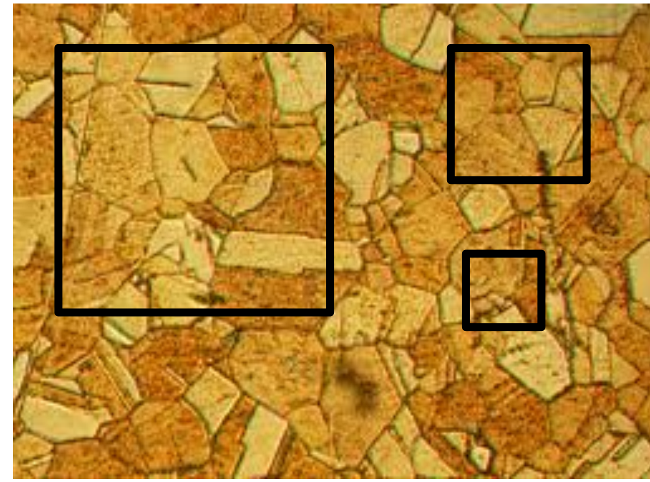


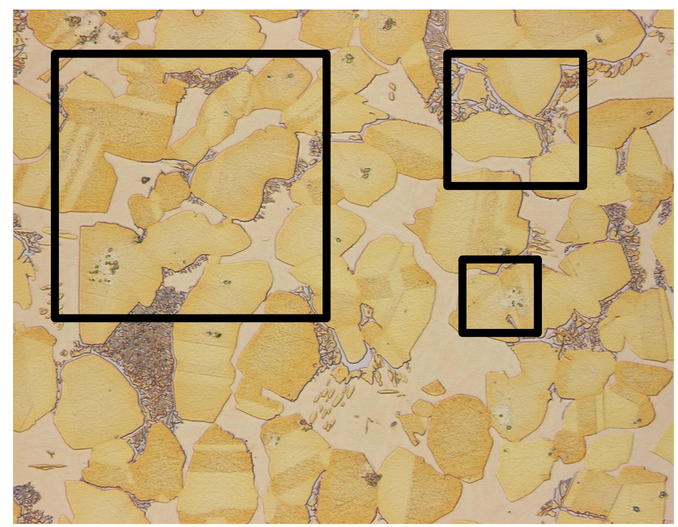
“As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality.”
— Albert Einstein

Our whole simulation approach ignores the primary source of aleatory uncertainty

- Most simulations (mechanics, heat transfer, shock, CFD, …) describes materials in a homogenous manner
- Materials are "painted" in to regions
- This becomes inappropriate at some point under mesh refinement (heterogeneity becomes too large to ignore.
- Taking this point of view has several profound effects:
  - The problem is now different, potentially by alot
  - The equations need to change – need subgrid effects
  - The models need to change – they should be length scale dependent
  - Solutions now have an important random component

We paint materials into regions most of the time





The governing equations themselves are a problem in this regard – mean field approach

- Most equations simply evolve the mean of the field

$$\frac{\partial U}{\partial t}+\nabla\cdot F\left(U\right)=0 \quad \Rightarrow \quad \frac{\partial \bar{U}}{\partial t}+\nabla\cdot F\left(\bar{U}\right)=0$$

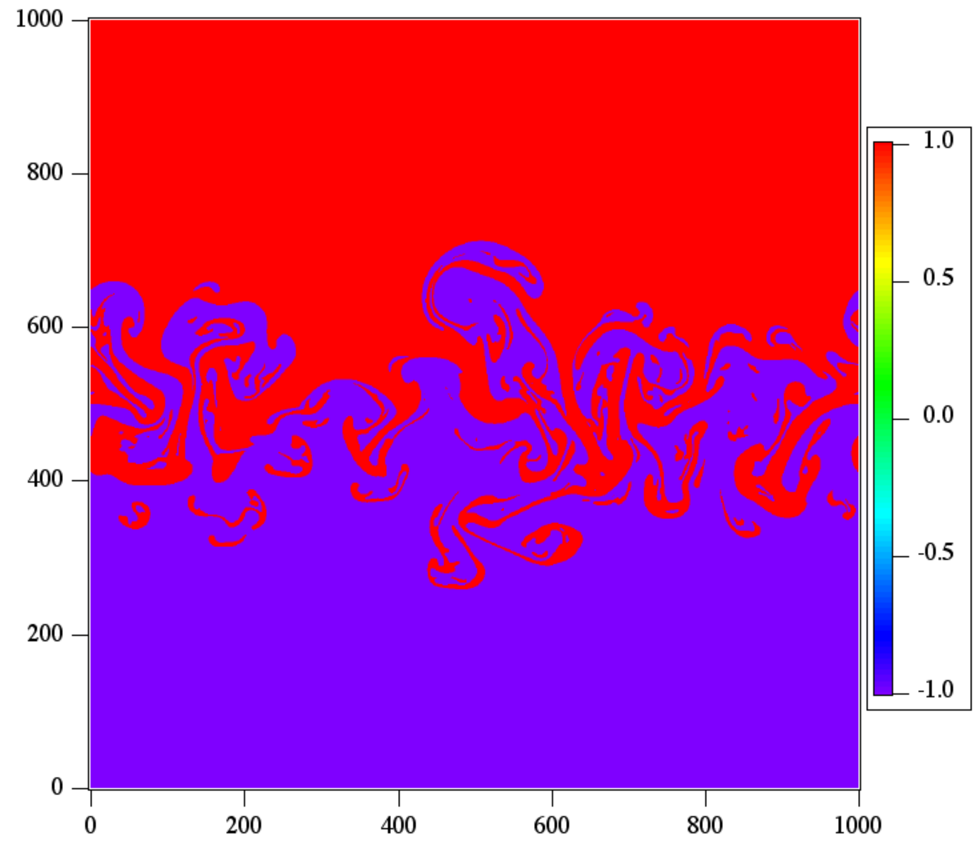
- One could start to take a philosophy more similar to LES in turbulence, and consider the nonlinear effect of fluctuations

$$\frac{\partial \bar{U}}{\partial t}+\overline{\nabla\bullet F\left(U\right)}=0 \quad \Rightarrow \quad \frac{\partial \bar{U}}{\partial t}+\nabla\bullet F\left(\bar{U}\right)+\nabla\bullet\tau\left(\bar{U}\right)=0$$

- The question is then the structure of the subgrid stress both the physics, but also the randomness of the field (the unresolved or unresolvable portion). **This change renders the solution non-deterministic**.

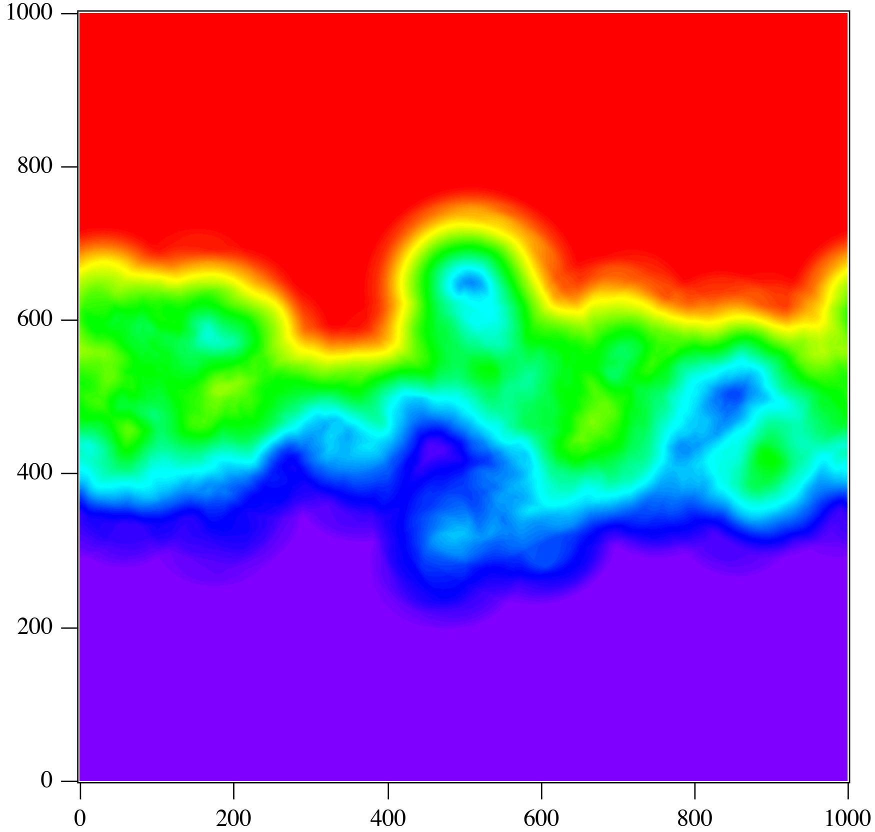
How to think about averaging

- Created by Tim Clark (formerly LANL, now U. New Mexico)
- How Rayleigh-Taylor mixing is viewed by RANS vs LES
- Emblematic of mean field modeling versus detailed modeling

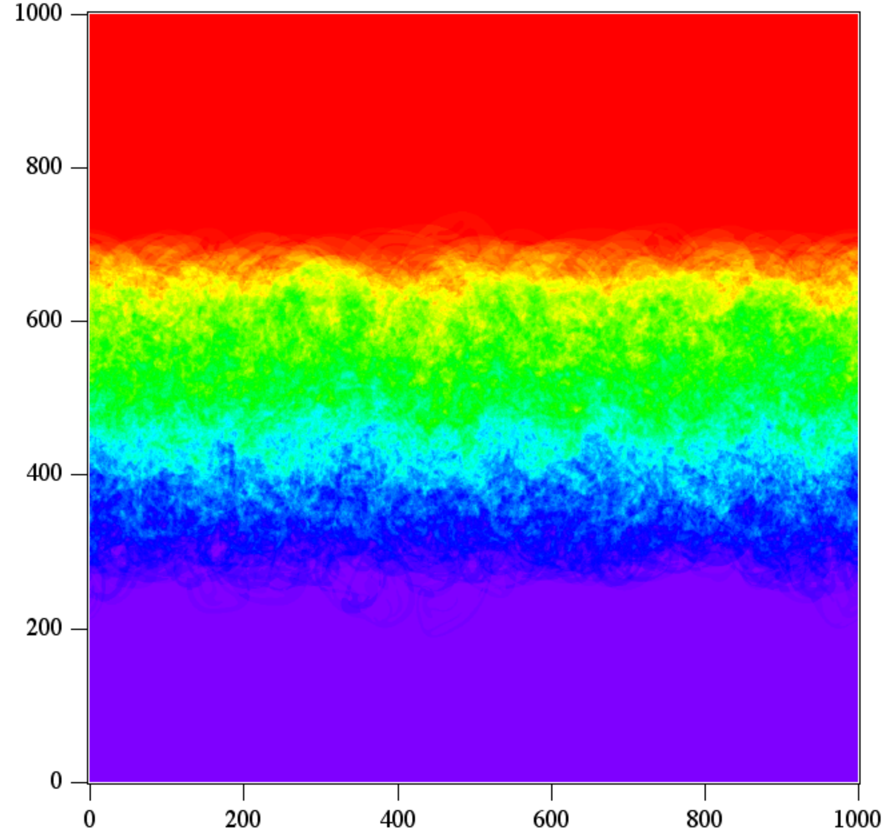


One Realization

Example of an Exact Ensemble Average



Radius = 64.5 “Cells”



Seventy Five Realizations

What is a RANS or LES Approach? (from Tim Clark, UNM). Current modeling is RANS-like

- Reynolds averaged Navier-Stokes approach implies Reynolds averaging, e.g.

$$u_i^{(n)} = \langle u_i \rangle + u_i'^{(n)} \qquad \langle u_i \rangle = \frac{1}{N_R}\sum_{n=1}^{N_R} u_i^{(n)} \qquad u_i' = u_i'^{(n)} \qquad \langle u_i' \rangle = 0$$

- The averaging is most conveniently considered to be an ensemble average.
- Models in this category are extremely general…
- LES "Averaging" - Spatial average

$$\bar{u}_i(\mathbf{x},t) = \frac{\int_{\Omega_r} f(\mathbf{r})u_i(\mathbf{x}+\mathbf{r},t)d\mathbf{r}}{\int_{\Omega_r} f(\mathbf{r})d\mathbf{r}}$$

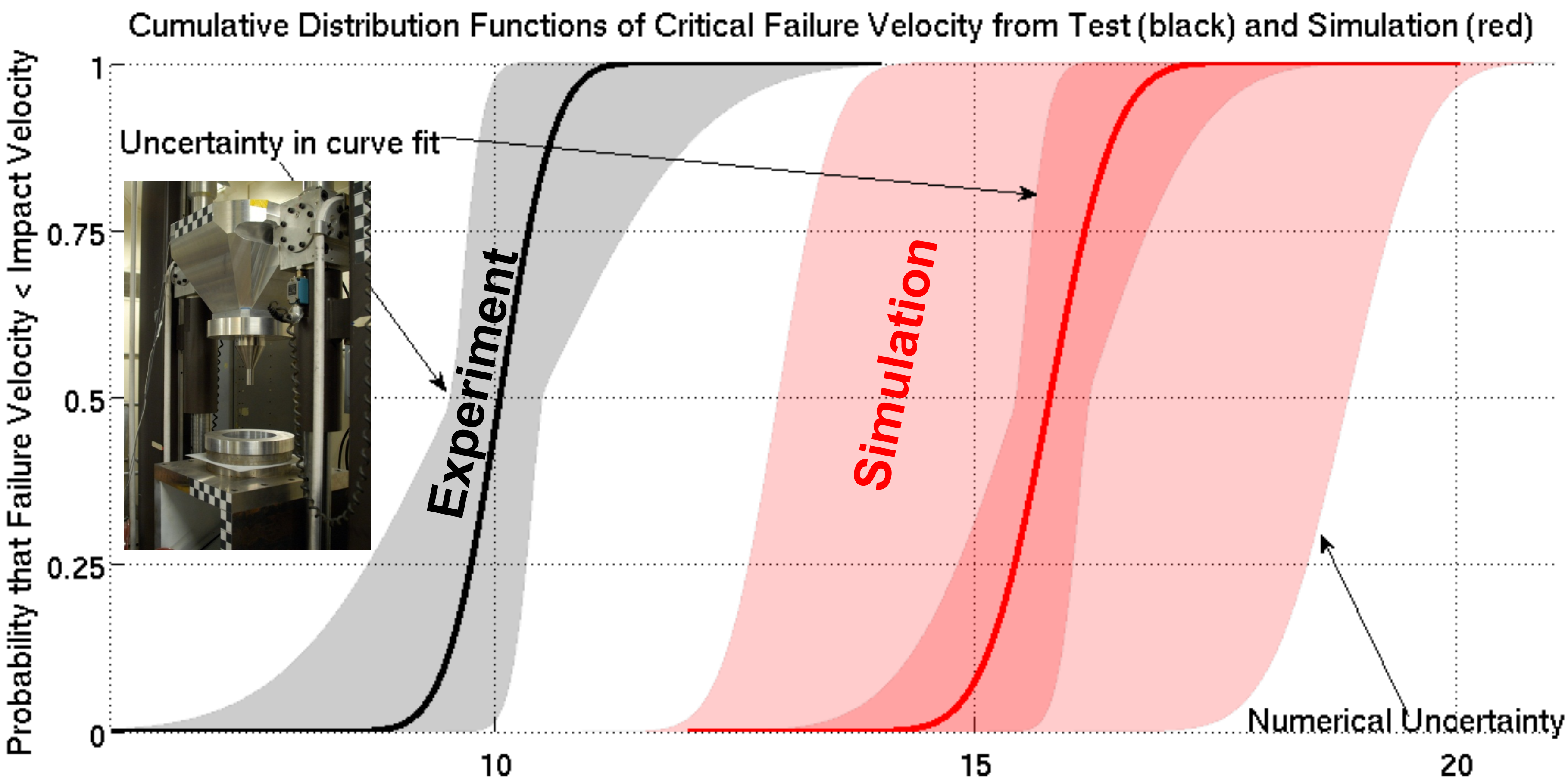
- Filtering leads to equation for the "filtered" (resolved) field and the subgrid field.
- Interactions include resolved-resolved fields, unresolved-resolved fields, and unresolved-unresolved fields.
- Models assume that subgrid scales are subservient to larger-scales.
  - Typically neglect transfer from unresolved field to resolved field.

“The mistake is thinking that there can be an antidote to the uncertainty.”
— David Levithan

Numerical error is too often ignored

- Quite often a simple mesh sensitivity analysis just declares “numerical errors are small”
  - Then fail to include it in the overall uncertainty.
  - Just including the numerical error at all is a huge step forward
- Are we putting the numerical error in the “right” place?
  - Is it the right size and sign
  - Is it biased properly and is that bias’ impact well defined.

Uncertainties in the Prediction and Experiment Comparisons



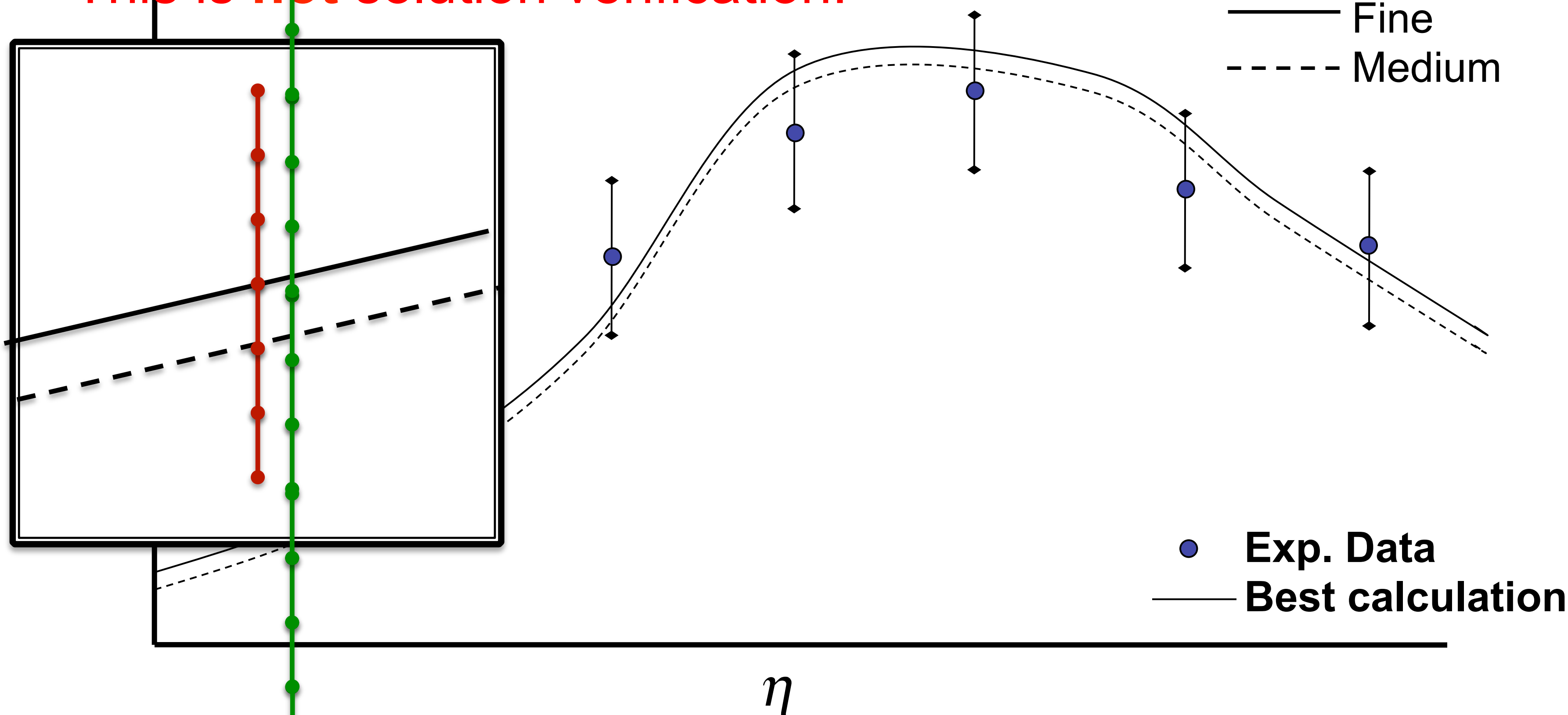
ainty in Experimental Impact Velocity = +/- 0.15 ft/s <<< uncertainty in fit

Numerical error is most often ignored by making optimistic assumptions

- Doing mesh sensitivity (sometimes or usually with two grids) then assuming things are OK because the change in solution is small.
- If the convergence rate is small enough, the size of the numerical error will be much larger than the difference especially if the grid is small
- For example, h1=4, h2=2, u1=500, u2=510 and convergence rate is actually 0.25
  - Converged value of "u" is 562.8, U_{num}=188
  - The standard CGI error is 30, a factor of six too small.

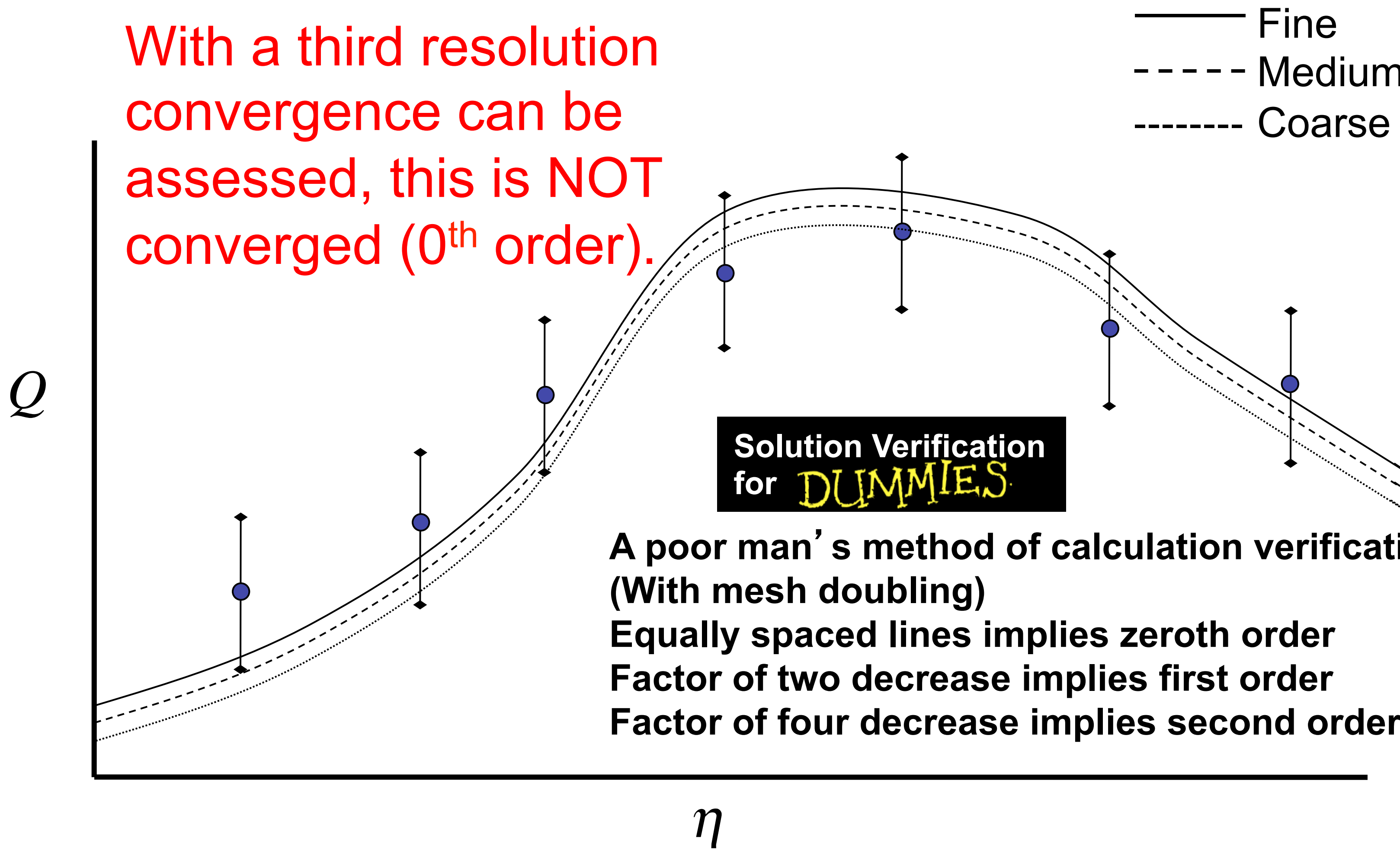
Here is a notion of how a "converged" solution might be described.

You might see this although rarely depicted in this manner.
This is **not** solution verification!



Here is a notion of how a "converged" solution might be described.

With a third resolution convergence can be assessed, this is NOT converged (0^{th} order).

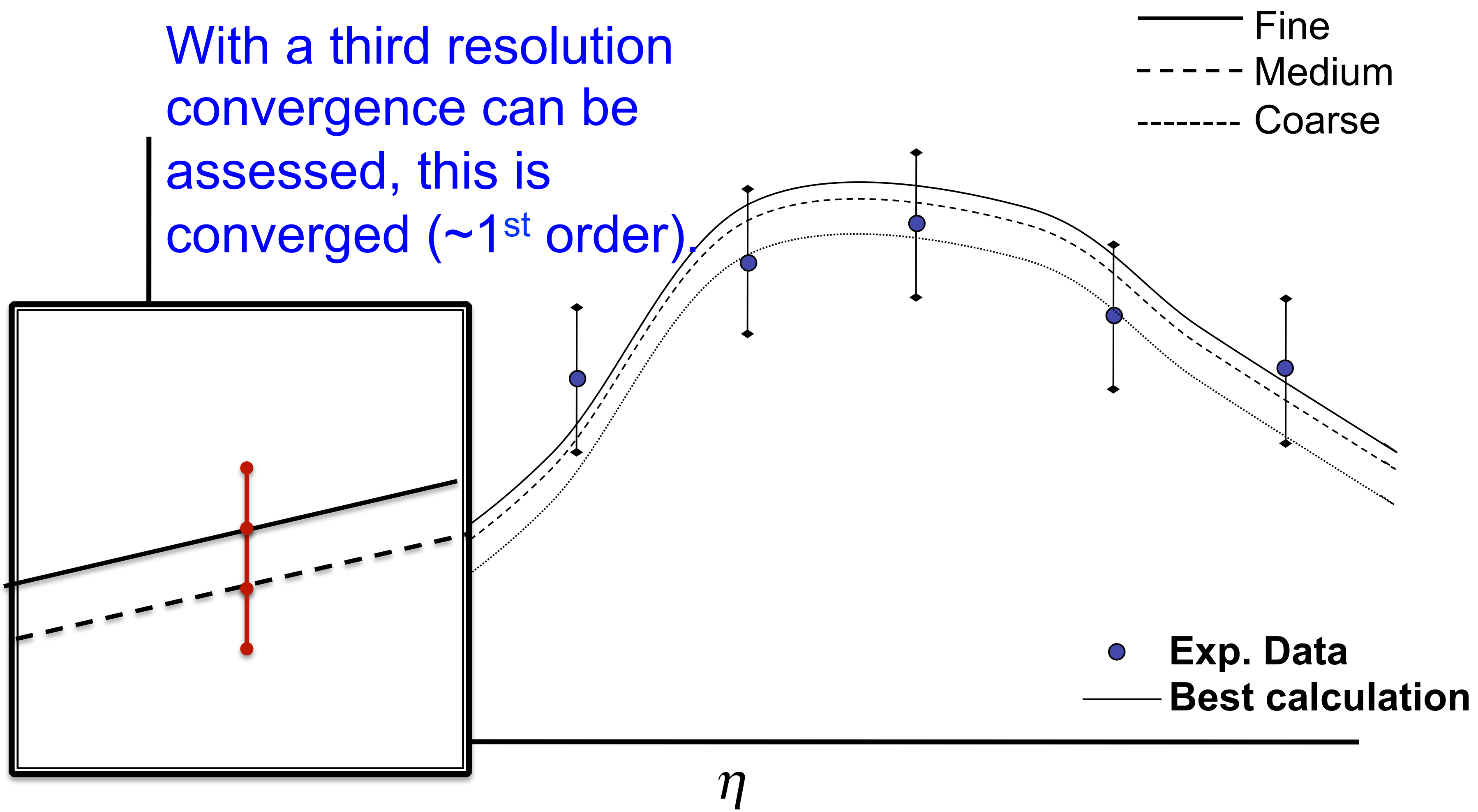


Solution Verification for DUMMIES

A poor man's method of calculation verification:
(With mesh doubling)
Equally spaced lines implies zeroth order
Factor of two decrease implies first order
Factor of four decrease implies second order

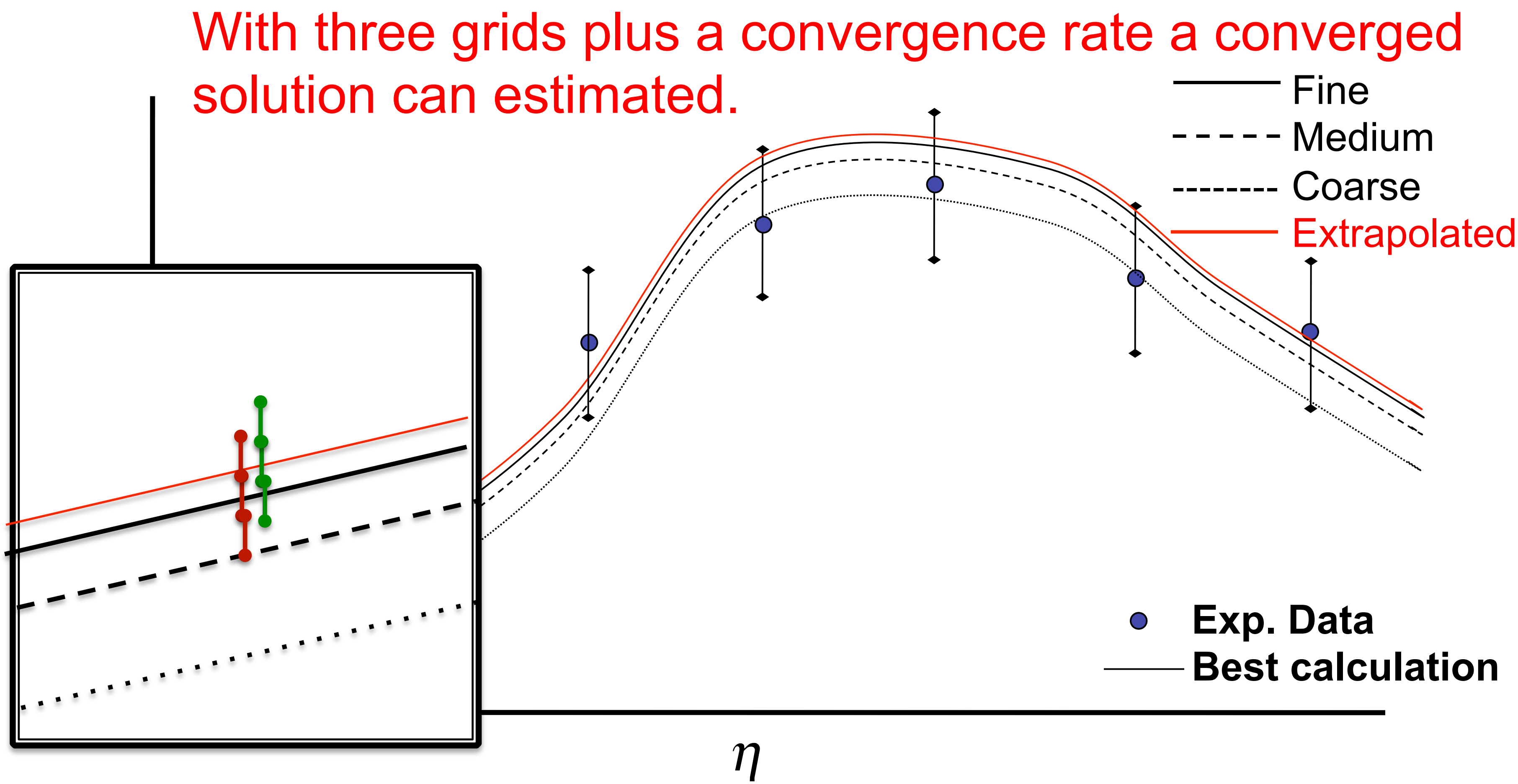
This is solution verification despite the bad results

Here is a notion of how a “converged” solution might be described.

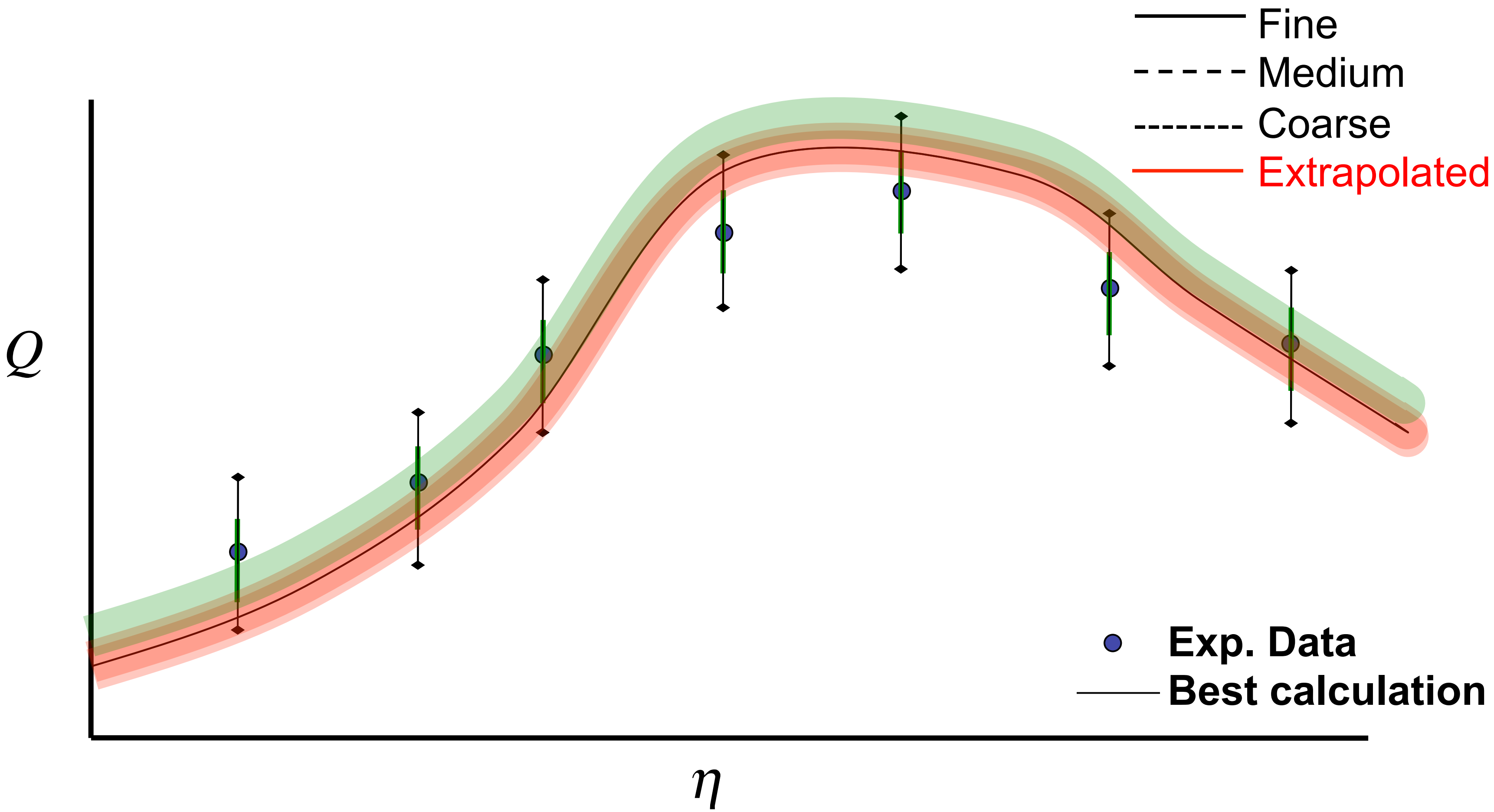


It is absolutely essential that the quantity “Q” be something that can reasonably be expected to converge numerically.

This sequence of meshes can be used to extrapolate the solution.



From this information a numerical error bar can be derived with rigor.



The standard setting for calculation verification.

Prefactor

Convergence rate

Mesh converged solution, i.e., $h \rightarrow 0$

$$\tilde{u} = u_h + A h^p$$

Characteristic length scale of mesh cell, "Δx"

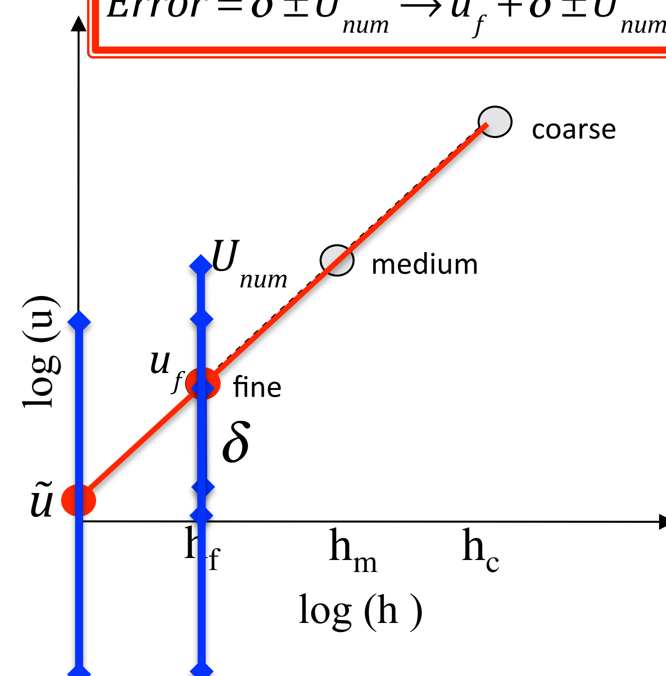
Computed solution on grid of mesh size h

$$u_{best} = \tilde{u} \pm U_{num}$$

$$Error = \delta \pm U_{num} \rightarrow u_f + \delta \pm U_{num}$$

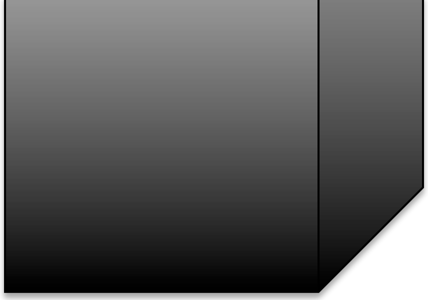
$$\delta = \tilde{u} - u_f \qquad U_{num} = F_s \left|\delta\right|; F_s = 1.25$$

- Where should the error bar be placed (i.e., centered)?
  - We have choices (two examined here):
    - Around the finest grid solution
    - Around the mesh converged solution
    - **The mesh converged solution is a best estimate and should define error.**
    - Error on the fine grid "should" be asymmetric.
    - The difference is significant



“Although our intellect always longs for clarity and certainty, our nature often finds uncertainty fascinating.”
— Carl von Clausewitz

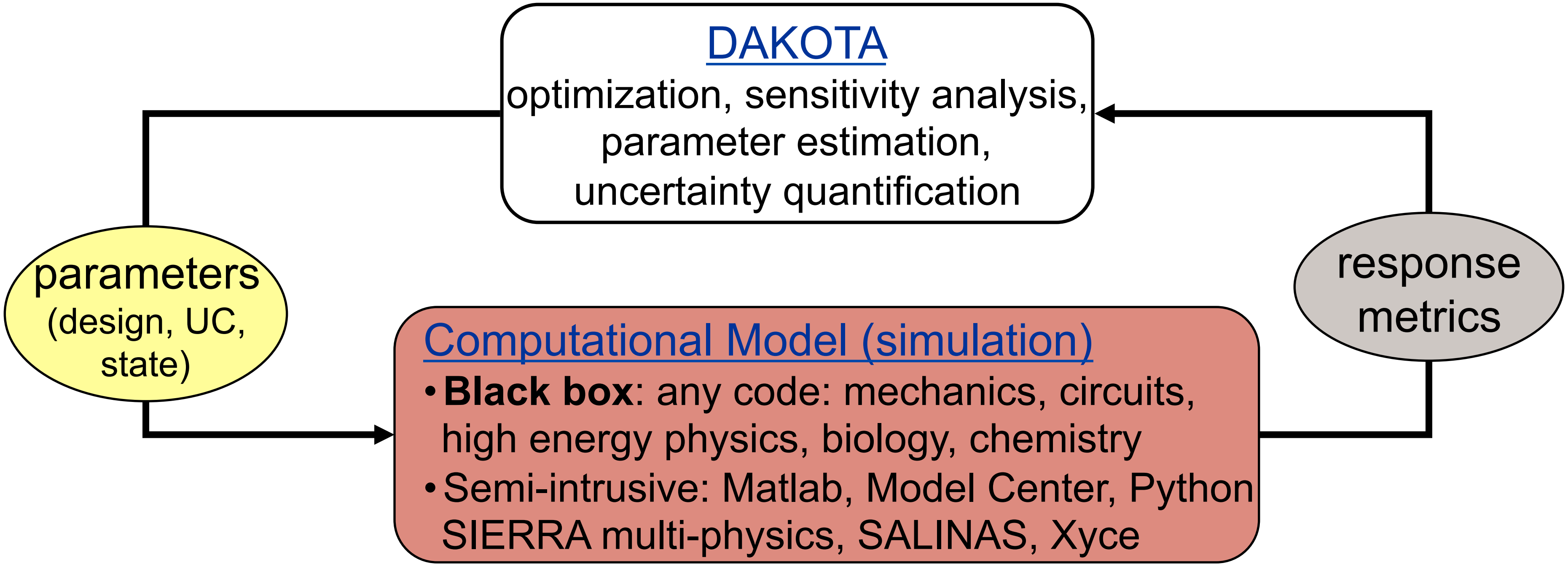
The "black box" mentality is hurting UQ



- Viewing codes and UQ in a black box manner leads to poor technical quality.
  - The best analysis is done by people who understand the inner workings of the codes they use.
  - The code is more of a "clear box" to them
- Take verification as an example: only by knowing and understanding the numerical methods in the code, and the problem being solved can results be understood properly.
  - Without such knowledge work is conducted naively
  - Modeling is susceptible to similar problems.

DAKOTA Provides an Automated Approach to Generate an Ensemble of Simulation Code "Runs"

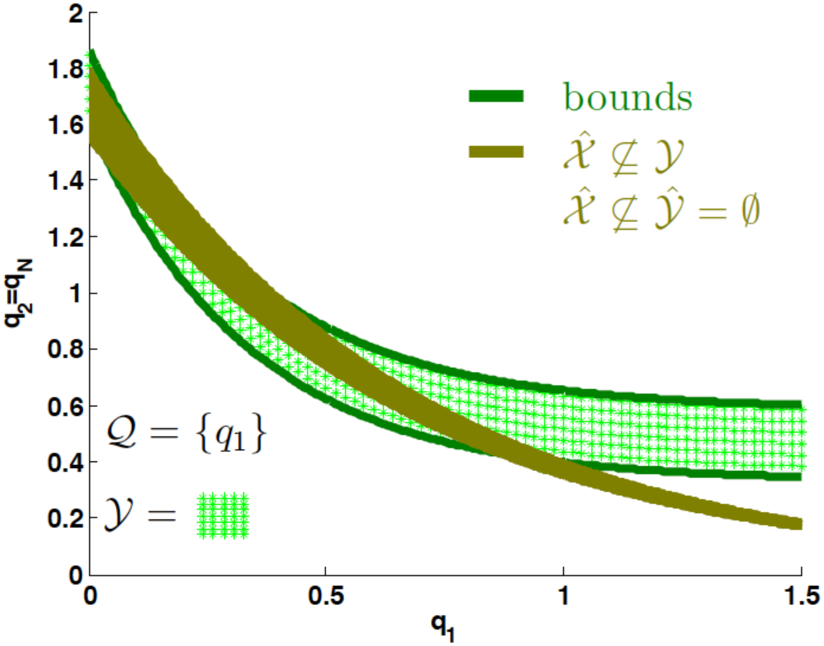
Automate typical "parameter variation" studies with a generic interface and advanced methods



- Can support experimental testing: examine many scenarios/ conditions with computer models, then physically test a few worst-case conditions.

Our codes should be “clear boxes”

- Commercial CFD and mechanics codes are accelerating this trend and allowing analysis to be done easily.
- Results can be impacted by this lack of knowledge
- Consider the example from Vaughan & Preston’s, “Physical Uncertainty Bounds”



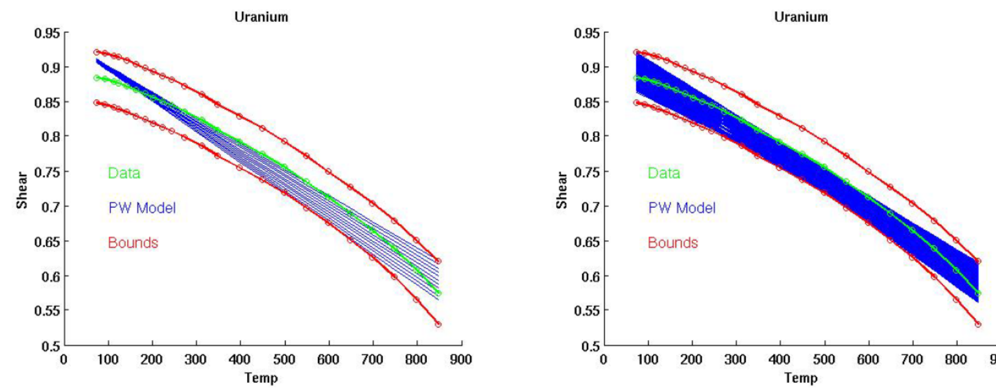


Figure 16: Uranium shear modulus versus temperature at $P = 0$; all curves as in previous figure. On the left, variations in the PW parameter α is shown as a family of blue lines. On the right we have possible variations that treat $G(\rho_0, T = 0)$ as a parameter, which implies no density dependence.

“Regardless of your faith, you can never escape uncertainty.”
— Shannon L. Alder

“The future is certain. It is just not known.”
— Johnny Rich

User error is unacknowledged and ignored

- Generally speaking we do not have a good idea of how different two solutions to the same complex problem can be based on who solved them
  - “The biggest determiner on the answer is the name of the analyst”
- We can guess that for very complex problems it can be very large, and potentially exceed other uncertainties.
- This issue is only strongly acknowledged by the nuclear safety community, some work is available in mechanics and CFD
- All point to the impact being significant to greater than other uncertainties.

Predictive Capability Maturity Model (PCMM)
Measures and Communicates Maturity of Mod/Sim Process

Sandia National Laboratories

PREDICTIVIE ATTRIBUTE	Level 0 Low-Consequence M&S-Informed, e.g., Scoping or Res Activities Score=0	Level 1 Low-Consequence M&S-Informed, e.g., Design Support Score=2	Level 2 High-Consequence M&S-Informed, e.g., Qualification Support, Score=4	Level 3 High-Consequence M&S-Based, e.g., Qualification Score=6
Representation or Geometry Fidelity Are you overlooking important effects because of defeaturing or stylization	• Grossly defeatured or stylized representation based on judgment or practical considerations	• Significant defeaturing or stylization based on judgment or practical considerations …	• Limited defeaturing or stylization judged to retain the essential elements of "as built" …	• Highest fidelity representation "as is" w/o sig defeaturing or stylization • or appropriate lower fidelity … justified w highest … ntation
Physics and Material Model Fidelity How science-based are the models?	• Un… wit… cal… • Em… sig… uni…			… physics-based model … nificant extrapolation
Code Verification Are software errors or algorithm deficiencies corrupting simulation results?	• Ju…			… and assessed … ainst SQE standards … ication test suite w … erage of required … interactions
Solution Verification Are numerical errors corrupting simulation results?	• Ju… • Se… alg… SR… de…	… context		… erical error bounds … RQs directly related … context
Validation How accurate are the models?	• Judgment only • Qualitative accuracy w/o significant SET coverage	• Qualitative accuracy w significant SET coverage • Quantitative accuracy w/o assessment of unc and w/o significant SET coverage	• Quantitative accuracy w/o assessment of unc • w significant SET coverage and IETs	• Quantitative accuracy w assessment of unc • w significant SET coverage, IETs, and full system test
UQ and Sensitivities What is the impact of variabilities and uncertainties on performance and margins?	• Judgment only • Deterministic assessment of margins (e.g., bounding analyses) • Informal "what if" assessments of unc, margins, and sensitivity	• Aleatory and epistemic uncertainties represented and propagated w/o distinction • Sensitivity to uncertainties explored	• Aleatory and/or epistemic uncertainties represented separately and propagated w significant strong assumptions • Quantitative sensitivity analysis w significant strong assumptions • Sensitivity to numerical errors explored	• Aleatory and/or epistemic uncertainties represented separately and propagated w/o significant strong assumptions • Quantitative sensitivity analysis w/o significant strong assumptions • Numerical errors quantified

Attempt to add a PCMM category related to code users was deemed "too hot to handle" and shelved.

Path forward: measured numerical error, white box codes, variability & user impact

- Changes are necessary to improve our uncertainty "coverage"
- Account for numerical error, always
  - Center the error properly and show any bias
- Account for the natural variability
  - Make the simulation like reality
- Test codes in a white box manner
  - See Preston and Vaughs paper
- Test the impact of different users (different modeling) on the results.

Combining uncertainty

- The "knee-jerk" approach is the RMS (L2) view,

$$U = \sqrt{U_{\exp}^2 + U_{epi}^2 + U_{ale}^2 + U_{num}^2 + U_{user}^2 + \ldots}$$

- This is fine as long as you have a good accounting of the sources and can justify the RMS point-of-view
- Other approaches: sum of absolute values (L1)

$$U = \left|U_{\exp}\right| + \left|U_{epi}\right| + \left|U_{ape}\right| + \left|U_{num}\right| + \left|U_{user}\right| \ldots$$

- Another approach: worst case (L infinity)

$$U = \max\left(\left|U_{\exp}\right|, \left|U_{epi}\right|, \left|U_{ale}\right|, \left|U_{num}\right|, \left|U_{user}\right|, \ldots\right)$$

“Maturity, one discovers, has everything to do with the acceptance of ‘not knowing.”

— Mark Z. Danielewski

Summary and Outlook

- A morphology of uncertainty is important to consider
- How realistic is an accurate, but pessimistic look at uncertainty.
- Our current approach is too epistemically focused, lazy and views codes as “black boxes”
- Codes need to be “white boxes”
- Numerical error needs to be taken seriously
- Aleatory uncertainty needs tools that focus on the sources of it
- Combining uncertainty should be done thoughtfully with respect to its aggregate effect.