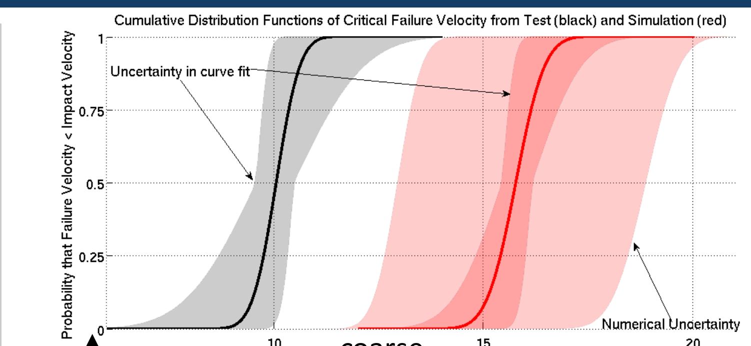
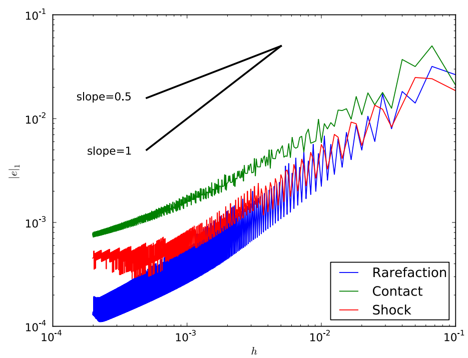
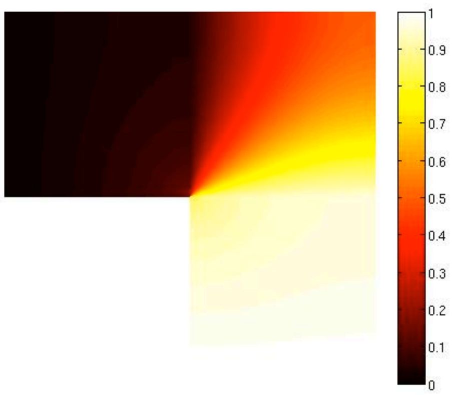
SAND2014-15361PE

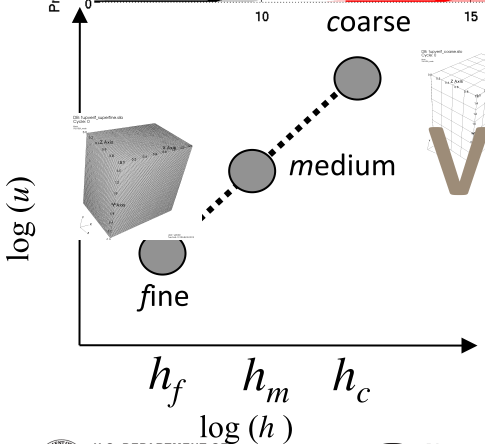
Exceptional service in the national interest

Sandia National Laboratories









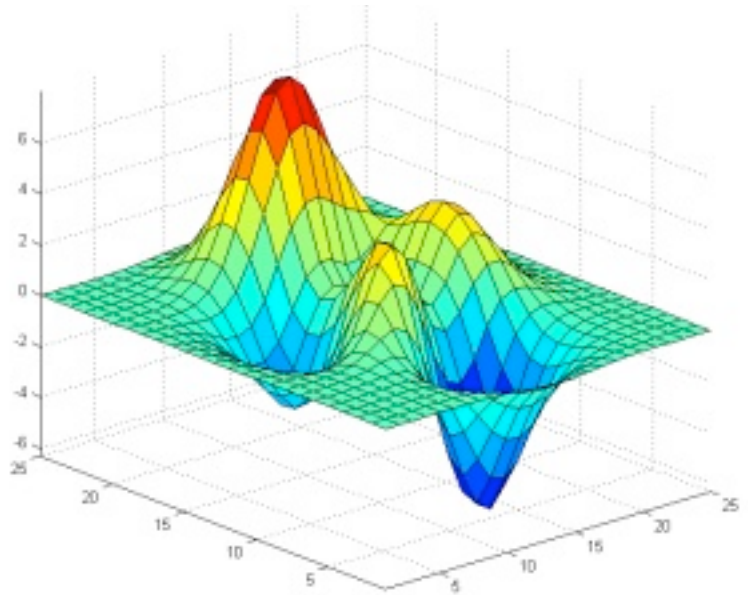
Verification and Model Selection

Bill Rider, WCCM, Barcelona July 21, 2014

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

What verification means in numerical analysis!

"For the numerical analyst there are two kinds of truth; the truth you can prove *and the truth you see when you compute*." – Ami Harten



Code Verification vs. *Solution* Verification

Code Verification:

- You have an exact solution, so you compute exact errors
- You are testing your code (implementation, algorithm)
- Hard estimates of convergence properties
- Metrics are defined by numerical analysis

Solution Verification:

- You don't have an exact solution, you estimate numerical errors
- You test your solution(s)
- Soft estimates of numerical error
- **Metrics are defined by the analyst** – integrated quantities, point values, functionals of the solution

The Standard Setting For Calculation Verification: Richardson extrapolation for error estimation

Sandia National Laboratories

- We begin with the standard error form,

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

- **Mesh converged solution, i.e., h→0** ($\tilde{u}$)
- Computed solution on mesh of char'c size h (u_h)
- **Prefactor** (A)
- Characteristic length scale of mesh cell, "Δx" (h)
- **Convergence rate** (p)

- The standard safety factor gives an uncertainty estimate (the *GCI**):

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

- This safety factor gives an ostensible 95% confidence interval,
  - ~2 std. dev. from CFD "experience" and computational experiments.
- Other forms will provide different estimates of F_s.
- For two grids, no estimate for δ is possible, and the uncertainty is intentionally "generous": $U_{num} = F_s \left|u_f - u_c\right| ; F_s = 3$

* GCI: Grid Convergence Index, i.e., Roache's approach

Error bars are subject to interpretation

Sandia National Laboratories

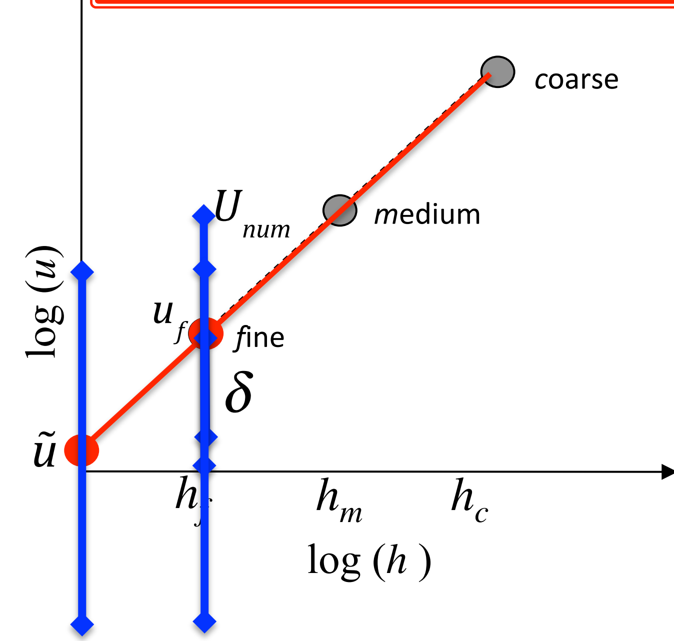
$$\delta = \tilde{u} - u_f \qquad\qquad U_{num} = F_s\left|\delta\right|; F_s = 1.25$$

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

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

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



There are some potential dangers to conscientiously avoid.

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

- This ansatz is valid for data in the asymptotic range of convergence.
  - Usually, we <u>assume</u> that the calculations are in the asymptotic range of convergence.
  - With two calculations, we have an under-determined fit through the results ( ).
  - With many calculations, the error ansatz is fully determined or over-determined; one can perform a **regression** fit.

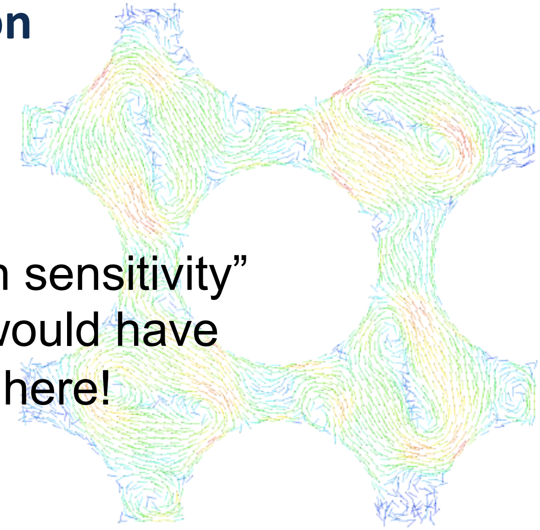
An example of how verification can go "off the rails"

Sandia National Laboratories

Preliminary Verification Results for CFD* for a CASL challenge problem (GTRF) with Fuego and Drekar (Δp), just spatial resolution

# ele.	Mesh	Fuego	Drekar
664K	Coarse	31.8 kPa	26.7 kPa
1224K	Medium	24.6 kPa	23.8 kPa
1934K	Fine	24.4 kPa	22.0 kPa

A "mesh sensitivity" study would have us stop here!



Fuego

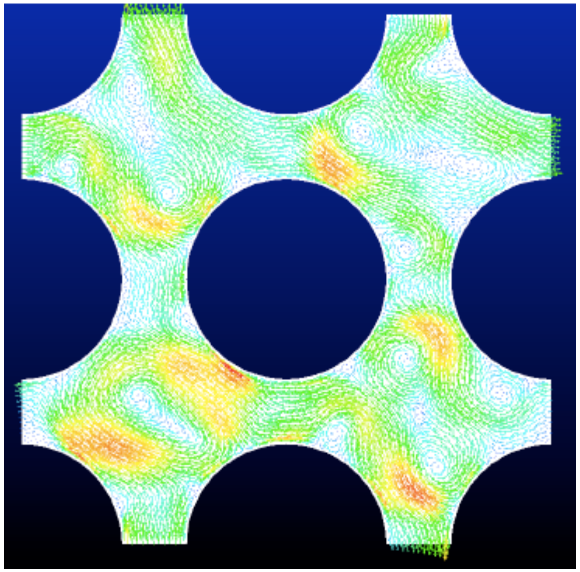
$$\Delta p(h) = 24340 + 26.6h$$

15.85!!

95 % Error Bound 80Pa (Roache, GCI) to 18.6kPa (Stern)

Drekar

$$\Delta p(h) = 17420 + 16370h^{1.234}$$



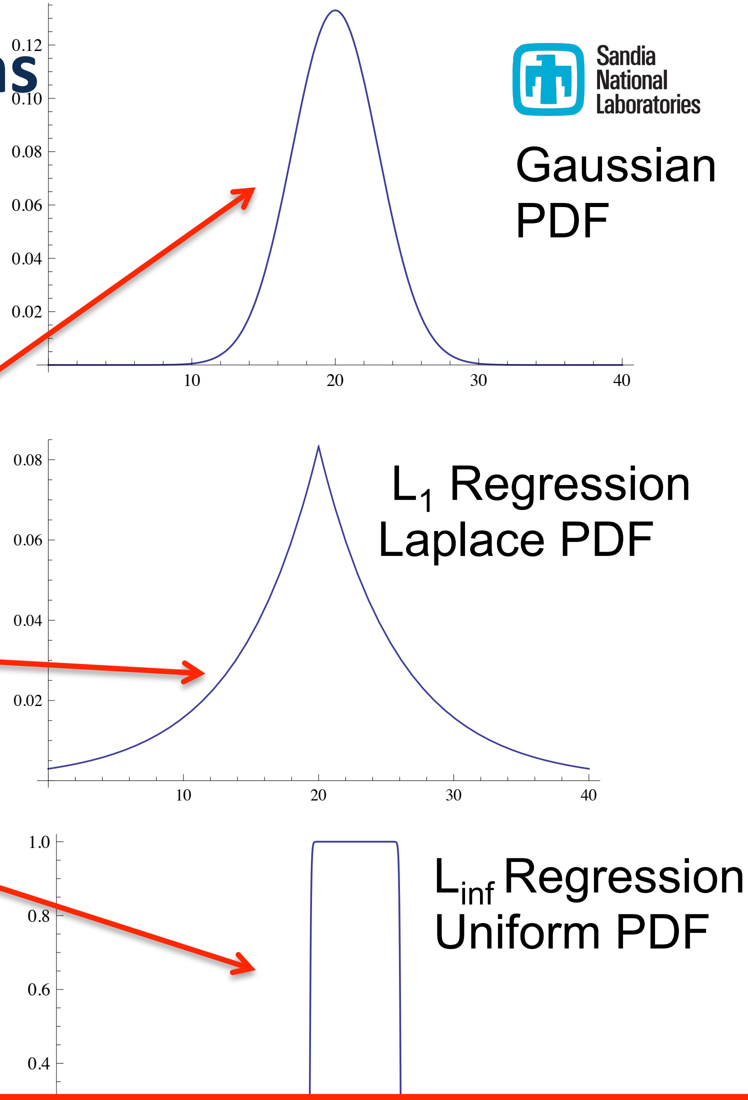
The Fuego result exemplifies one of the problems with the standard setting, 16th order convergence is absurd!

Preview: Our procedure gives *a Δp=16.1kPa ±13.5kPa.*

***CASL: Consortium for the Advanced Simulation of Light Water Reactors,** CFD: computational fluid dynamics
GTRF: Grid-to-rod-fretting

Brief Digression: Regression, norms and probability distributions

- Minimization of the residual for regression carries implications about optimality. The fit is optimal if the errors are distributed:
  - Gaussian implies L_2, *the standard approach* (unweighted)
  - Laplace (double exponential) implies L_1 (absolute value)
  - Uniform implies $L_{infinity}$ (maximum)
- *Regression can be done in any norm if the data is either under- or over-determined and can include constraints as well.*



Use robust statistics not standard statistics

We define a robust multi-regression (RMR) method to encode expert information, with robust statistics.

Sandia National Laboratories

- **The standard method is fragile and includes implicit assumptions regarding statistics, error and convergence, which are not included in the analysis.**
- We apply the following algorithm to the data (2 or more grids):

1. **A structured principled way to introduce specific expert knowledge.** Bounds on extrapolated solution could be entered too (such as positivity).
2.
3. **Produce a set of estimates with defined explicit assumptions regarding statistics of the error, and free of the fragility of a single estimate.**
4. **Apply robust statistical techniques to produce results with confidence. Run over subsets of data (jack-knife).**
5. **Produce a bounding estimate using the same approach. Useful when the data is non-monotonic.**

We can demonstrate our method by solving a linear ODE.

- The equation is trivially solvable. We use a forward Euler method here.

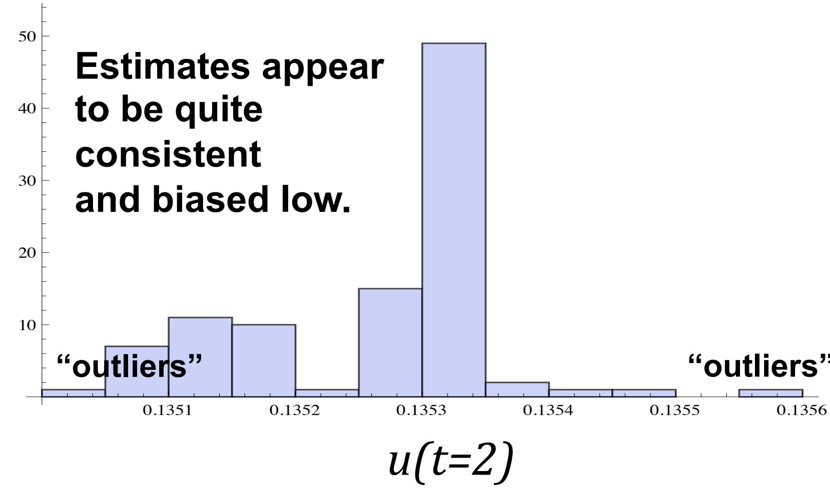
$$\dot{u} = -u \rightarrow u^{n+1} = u^n - \Delta t u^n$$

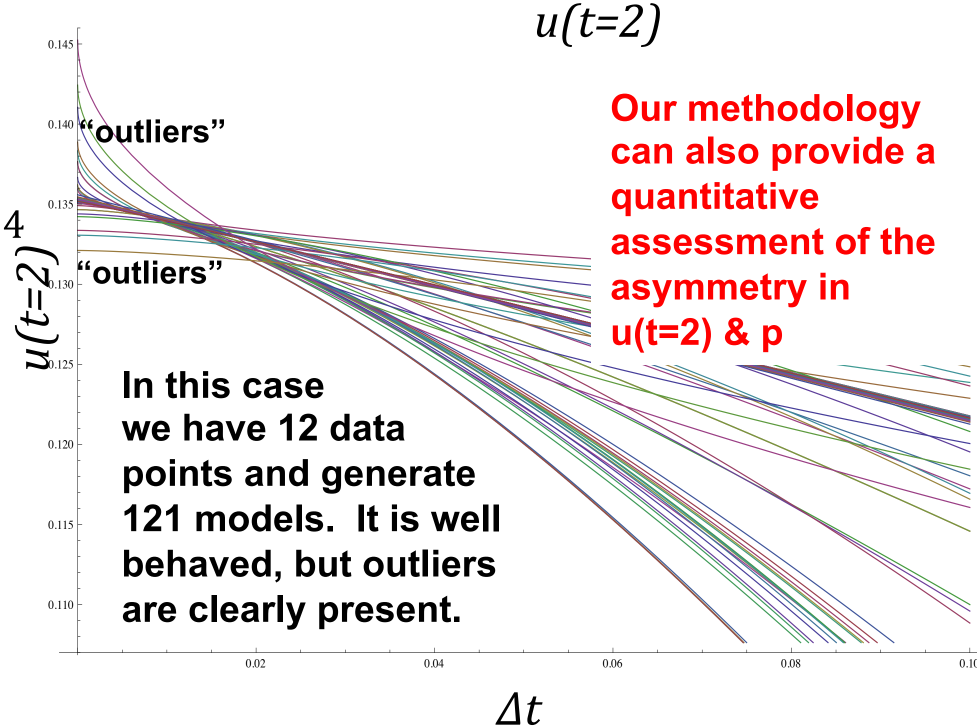
- The estimate and uncertainty strongly coincide with the analytical result.

$$u(t=2) = 0.135316 \pm 0.000138247, p = 1.0219 \pm 0.0154$$

  - The uncertainty "fans out" for Δt smaller than the finest time step used for computed results.
  - Roache's GCI approach does not capture the analytical solution!

$$u(t=2) = 0.134794 \pm 0.000517851, p = 1.03878$$





Solution Verification for Drekar gives estimates are very close to the three grid estimates given earlier!

Sandia National Laboratories

Mesh	Δp (pa)
671K	23400
1049K	26781
2664K	23804
5832K	22040
12522K	20745

RMR (all)

$$\Delta p = 16500 \pm 5435 + Ah^{0.53 \pm 0.09} Pa$$

RMR (drop coarse grid)

$$\Delta p = 16280 \pm 7060 + Ah^{1.08 \pm 0.27} Pa$$

GCI (all)

$$\Delta p = 23250 \pm 1635 + Ah^{0.5} Pa$$

GCI (drop coarse grid)

$$\Delta p = 16450 \pm 6960 + Ah^{1.05} Pa$$

- The results are convergent, but at a rate comparable with expectations for "rough" flows like turbulence, that is 1st order accurate.
- Uncertainties are large.
- RMR gives reliable results with or without dropping the non-convergent data point (671K)
  - The CGI is extremely unreliable without screening the data.

The estimated solution is frighteningly close to the estimates from the earlier Fuego calculations!

*RMR:Robust multi-regression, CGI: Grid Convergence Index

Next: Investigate Model Selection Procedures

- Verification is usually done with an assumed error ansatz
  - Almost always a power law form
- It would be beneficial to assume less, and define a better error model more supported by the data.
- Statistical estimation procedures exist for examining the form of the model that best fits the data
- Examples: AIC, BIC, **LASSO**
- Note, that these methods are well-defined for linear models, and verification has been shown to generally require nonlinear models

We decided to explore the LASSO Method in this context

- LASSO has already been applied to verification as one of the fits used in the robust verification work
- In its linear form it is a simple regularized least squares method

$$\min \left\| A\mathbf{x} - \mathbf{b} \right\|_2 + \lambda \left\| \mathbf{x} \right\|_1$$

  - As the Lagrange multiplier is increased in size the solution becomes increasingly "sparse".
  - Closely related to $\min \left\| \mathbf{x} \right\|_1$ constrained by $A\mathbf{x} = \mathbf{b}$
- **Our greatest leap is applying it to the nonlinear model selection**

$$\min \left\| f\left(\mathbf{x}\right) - \mathbf{b} \right\|_2 + \lambda \left\| \mathbf{x} \right\|_1$$

Variants of LASSO that may be better

- One does not have to do a regularized least squares.
- The L1 norm might be useful

$$\min \left\| f\left(\mathbf{x}\right) - \mathbf{b} \right\|_1 + \lambda \left\| \mathbf{x} \right\|_1$$

- …or the Danzig estimator using the infinity norm.

$$\min \left\| f\left(\mathbf{x}\right) - \mathbf{b} \right\|_\infty + \lambda \left\| \mathbf{x} \right\|_1$$

- All work the same way, as the Lagrange multiplier becomes large most of the coefficients go to zero.

The L1 Norm has some remarkable properties useful for model selection

Sandia National Laboratories

- *Some have said that L1 is "magic"*
- Specifically, the L1 norm promotes sparsity under the minimization framework
- This is used in compressed sensing (and basis pursuit)

$$\min \|\mathbf{x}\|_1 \text{ constrained by } A\mathbf{x} = \mathbf{b}$$

$$\min \|\mathbf{x}\|_1 \text{ constrained by } \|A\mathbf{x} - \mathbf{b}\|_2 < \varepsilon$$

- The approach allows us to rank the portions of the model form piece-by-piece by examining the terms that remain nonzero for a changing Lagrange multiplier.

“Any sufficiently advanced technology is indistinguishable from magic.”

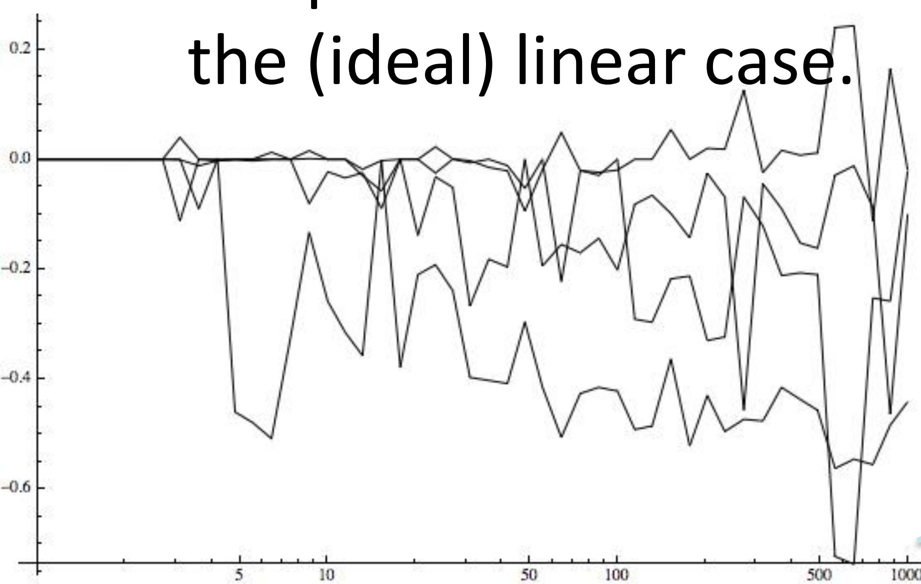
– Arthur C. Clarke

Simple example with model selection

- I will use the simple ODE problem as the first test.
- Broaden the potential terms in the model to include other possible terms: $\tilde{u} = u_h + A h^p$

$$\tilde{u} = u_h + A_p h^p + A_1 h + A_2 h^2 + A_e\left[1 - \exp(b_e h)\right] + A_L \log(1 + b_L h)$$

- The results are unexpected. The trends are clear, but the plots of the coefficients are not nearly as clean as the (ideal) linear case.



$$\min \left\| \tilde{u}(\mathbf{A}) - u_h \right\|_\infty + \lambda \left\| \mathbf{A} \right\|_1$$

$$\tilde{u} = u_h + A_1 h + A_2 h^2$$

Let's look at the performance as error estimators term-by-term

- We use a L1 estimator for each (if you do one fit, I'd recommend L1 instead of L2!)

$$u_{\text{true}} = \exp(-2) = 0.135335$$

- The linear plus quadratic was chosen (polynomial)

$$u_{1,2} = 0.135337 - 0.135337h - 0.0211486h^2; r = 0.0000609$$

- Exponential

$$u_{\text{exp}} = 0.135337 - 0.454392\left[1 - \exp(0.298264h)\right]; r = 0.000146$$

- Logarithmic

$$u_{\log} = 0.135470 - 2.130270\log(1 + 0.298264h); r = 0.002867$$

- Power Law $u_{\text{power}} = 0.135235 - 0.148584h^{1.03656}; r = 0.000640$

Example with picking out cross terms

- We are testing a model with two-dimensional transient heat conduction and chemistry for solute deposition in nuclear reactor cores. The code is poorly documented and it is not clear how coupled it actually is.
- We are going to use the LASSO to investigate what the error model should look like

$$\tilde{u} = u_h + A_x \Delta x^{p_x} + A_y \Delta y^{p_y} + A_t \Delta t^{p_t} + A_{xy} (\Delta x \Delta y)^{p_{xy}} + A_{xt} (\Delta x \Delta t)^{p_{xt}} + A_{yt} (\Delta y \Delta t)^{p_{yt}} + A_{xyt} (\Delta x \Delta y \Delta t)^{p_{xyt}}$$

  - If we just use every term over-fitting is a distinct possibility
- We find that the model that stands out is rather different than we would have chosen a priori. This model does well,

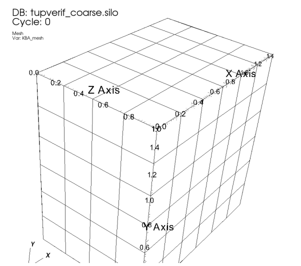
$$\tilde{u} = u_h + A_x \Delta x^{p_x} + A_y \Delta y^{p_y} + A_t \Delta t^{p_t} + A_{xy} (\Delta x \Delta y)^{p_{xy}} + A_{xyt} (\Delta x \Delta y \Delta t)^{p_{xyt}}$$

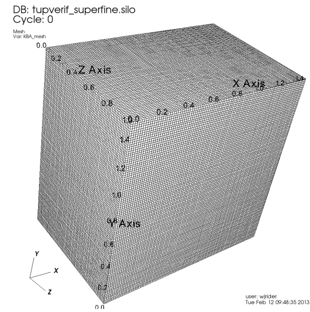
Applying this to Sn Neutronics

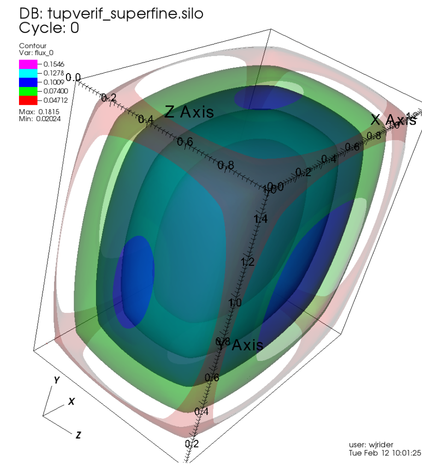
- We apply the verification techniques to the ORNL neutronics code Denovo, which solves the transport equation using discrete ordinates.

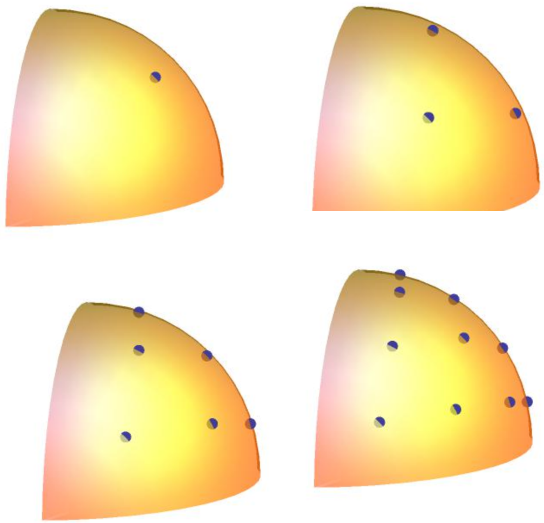
$$\Omega\cdot\nabla\psi(x,\Omega,E)+\sigma(x,E)\psi(x,\Omega,E)=\int dE'\int_{4\pi}d\Omega'\sigma_s(x,\Omega'\cdot\Omega,E'\to E)\psi(x,\Omega',E')$$

- The solution depends on six “coordinates” three space, two angular, and energy.
- We reduce this to space and angle while examining their distinct discretization.
- At this point only very simple problems have been examined.









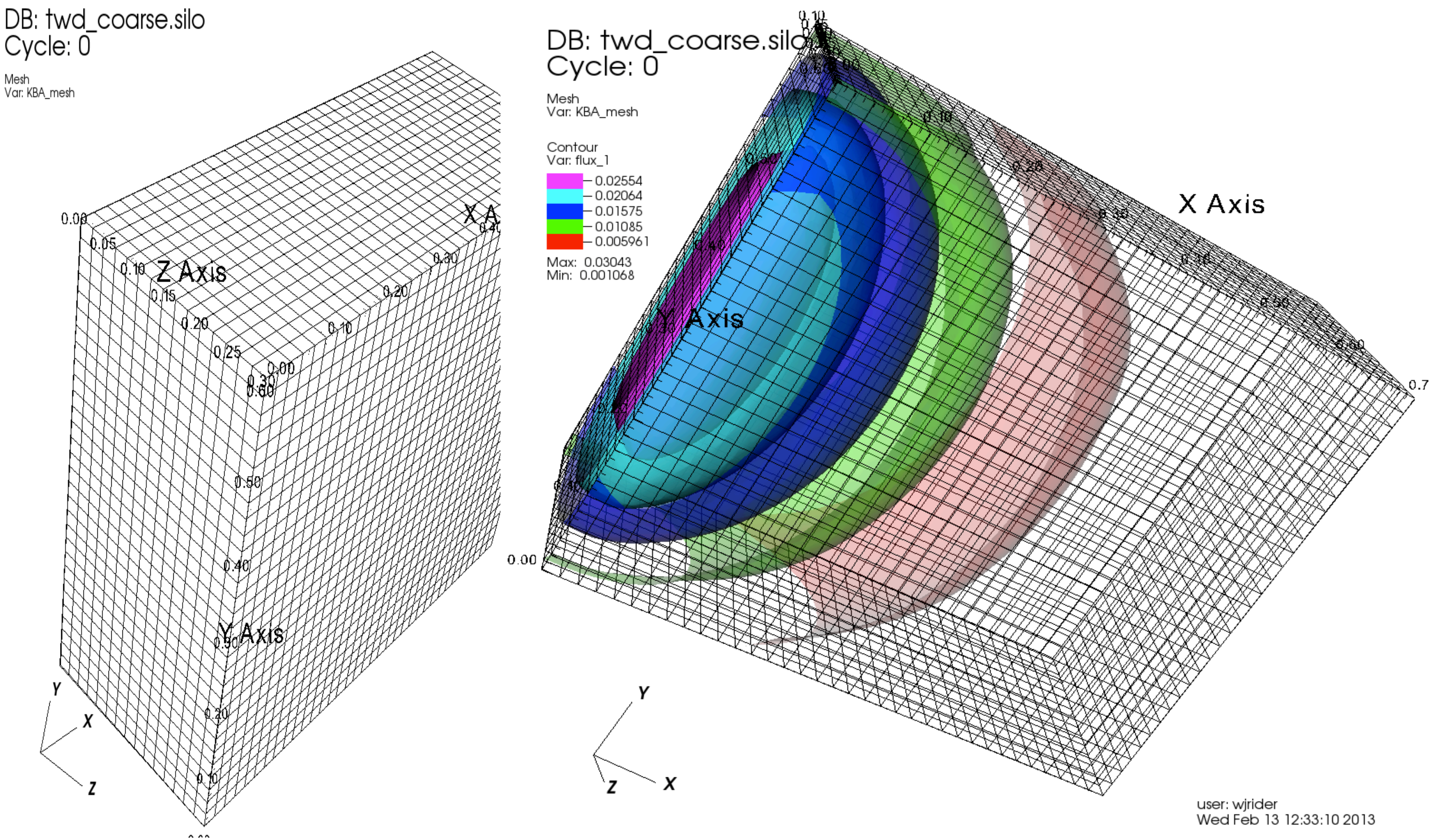
We need to modify the standard verification setting to accommodate the discrete ordinates method

- The basic approach remains, but we have two variables that convergence critically depends upon,

$$\tilde{u} = u_h + A h^p + B / n^q + C \left(h / n \right)^r$$

- A secondary issue is the theoretical expectations for the convergence with respect to the number of quadrature points where we have chosen a form like Jarrell (2010),
- We have examined four different spatial discretizations of the streaming term: Step characteristic, linear discontinuous, trilinear discontinous, and diamond differencing with negative flux fix-up
- We have examined four different quadratures: level symmetric, Gauss-Legendre, Quadruple Range, and LDFE.

The integrated scalar flux for the downscatter problem.



Neutronics Results

- We find little support for the h-p version of the error model
  - The third term in the error model drops out immediately as the size of the Lagrange multiplier is increased.
- We also examine the nature of the error model itself as before

$$\tilde{u} = u_h + \boxed{A h^p} + \alpha\left[1-\exp(\kappa h)\right] + B/n^q + \beta\left[1-\exp(\eta/n)\right] + \boxed{C(h/n)^r} + \gamma\left[1-\exp(\xi h/n)\right]$$

- We find the power law in space and the coupled space-angle power term dominate.

Summary and Future Work

- The basic idea appears to work well.
  - We have concerns about applying this to nonlinear models where theory is weaker.
- We can use it to determine the dominant terms in a proposed error model
- The exponential terms are repeatedly chosen by the procedure, this needs further examination
  - The exponential terms are a better basic error model
  - The LASSO procedure is somehow biased toward them
- We will continue the exploration.

Who Am I ?

- I'm a staff member at Sandia, and I've been there SNL for 7 1/2 years. Prior to that I was at LANL for 18 years. I've worked in computational physics since 1992.
- In addition, I have expertise in hydrodynamics (incompressible to shock), numerical analysis, interface tracking, turbulence modeling, nonlinear coupled physics modeling, nuclear engineering…
- I've written two books and lots of papers on these, and other topics.