SAND 2012-2563C

Cross-Application of Uncertainty Quantification and Code Verification Techniques

SIAM Conference on Uncertainty Quantification

April 3rd, 2012

V. Gregory Weirs
Sandia National Laboratories

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.

Acknowledgements

Some of this work is the subject of a talk given at the SAMO Conference, which became an article in RESS (accepted for publication.): Co-authors:

- Jim Kamm, Laura Swiler, Brian Adams, Bill Rider, Mike Eldred at SNL
- Stefano Tarantola and Marco Ratto from the Joint Research Center, European Commission, Ispra (VA), Italy

In this talk I'll focus more about the problem formulation from the RESS article than the results.

Sandia National Laboratories

Overview

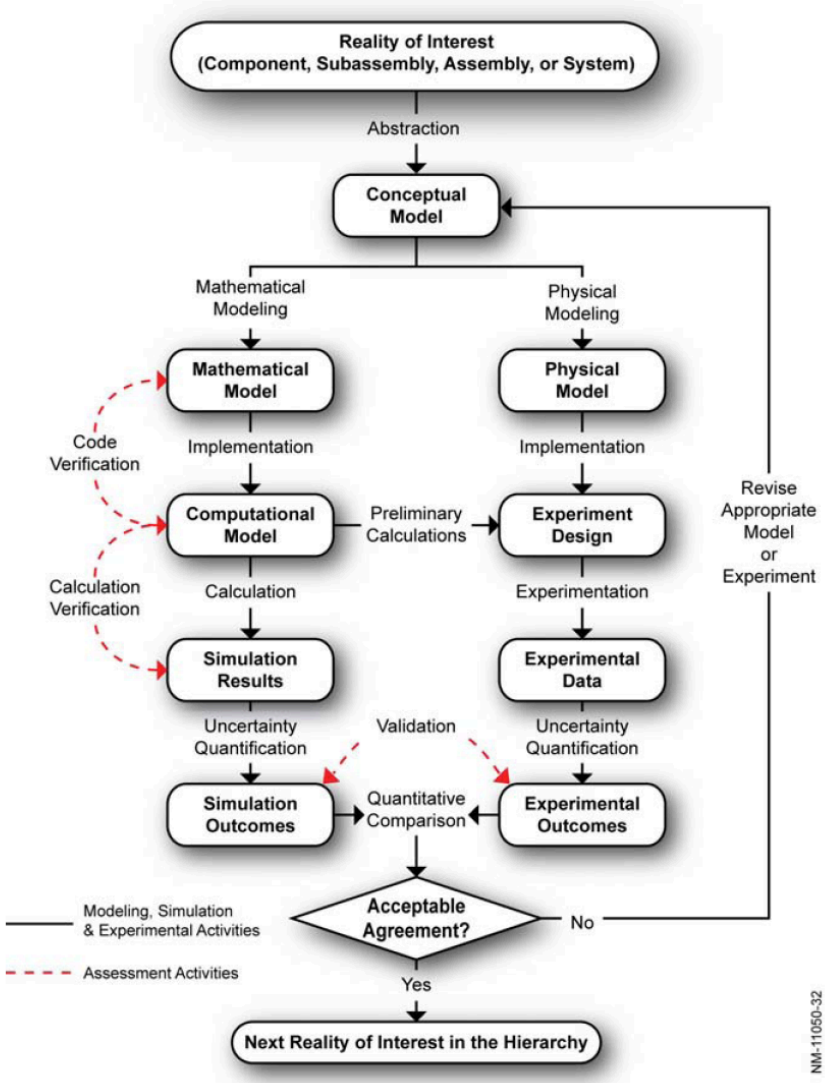
Each of the following techniques are useful for a variety of things, but consider these informal definitions for the present context:

- Uncertainty quantification is determining the probability distribution of an output given a model and probability distributions of uncertain inputs
- Sensitivity analysis is UQ with attribution
- Code verification is learning about your numerical methods through careful numerical experiments and exact solutions

How can we use these together? Is that even possible?

Sandia National Laboratories

Overview



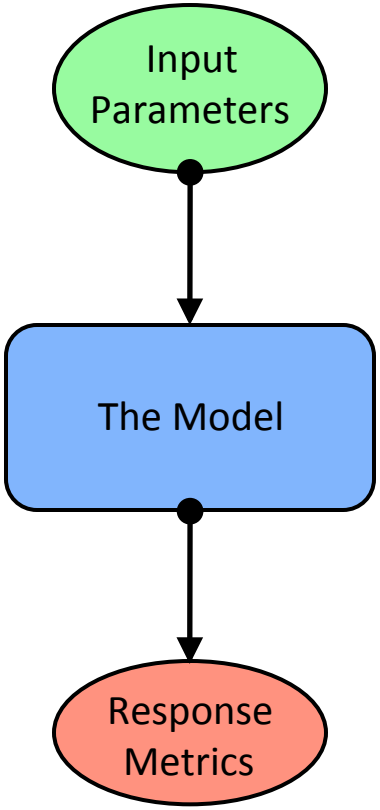
- This V&V process flowchart is taken from the ASME Solid Mechanics V&V guide.
- Note the positions of code verification, calculation verification, validation, and UQ in this workflow diagram.

ASME, *V&V 10-2006 Guide for Verification and Validation in Computational Solid Mechanics*, American Society of Mechanical Engineers (2006).

The Sensitivity Analysis Story

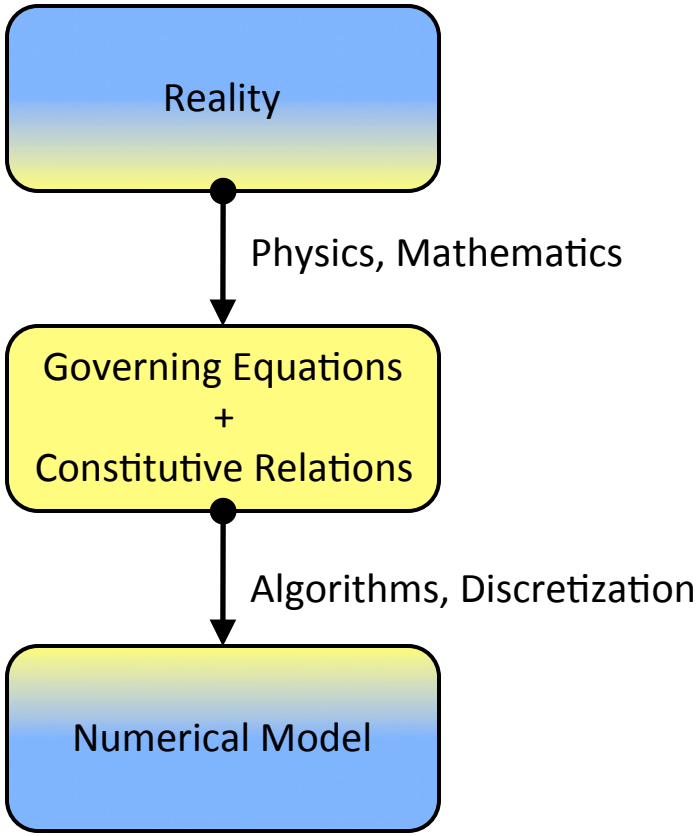
- Our problem of interest has 4 inputs and 4 outputs
- Outputs (responses) have different properties:
  - monotonic vs. non-monotonic
  - smooth vs. discontinuous
  - noisy vs. clean
- We examine different SA techniques:

  LHS, LP-Tau } sampling
  PCE } stochastic expansion
  SDP, ACOSSO, DACE } surrogates
- We compute sensitivity indices and compare them to exact values; in particular, we examine performance with respect to sampling resolution

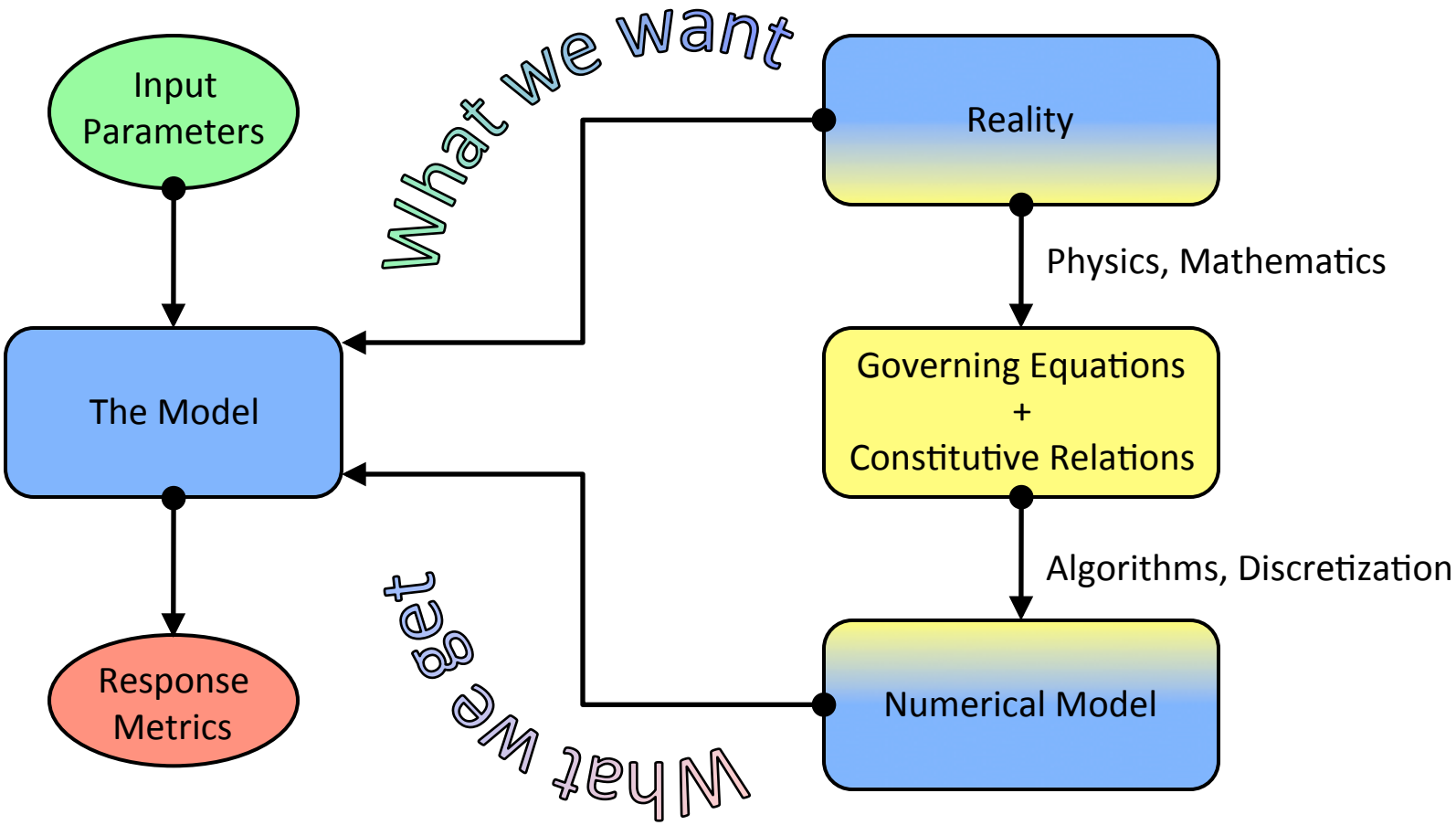


Verification and Validation
The ~~Application Space~~ Story

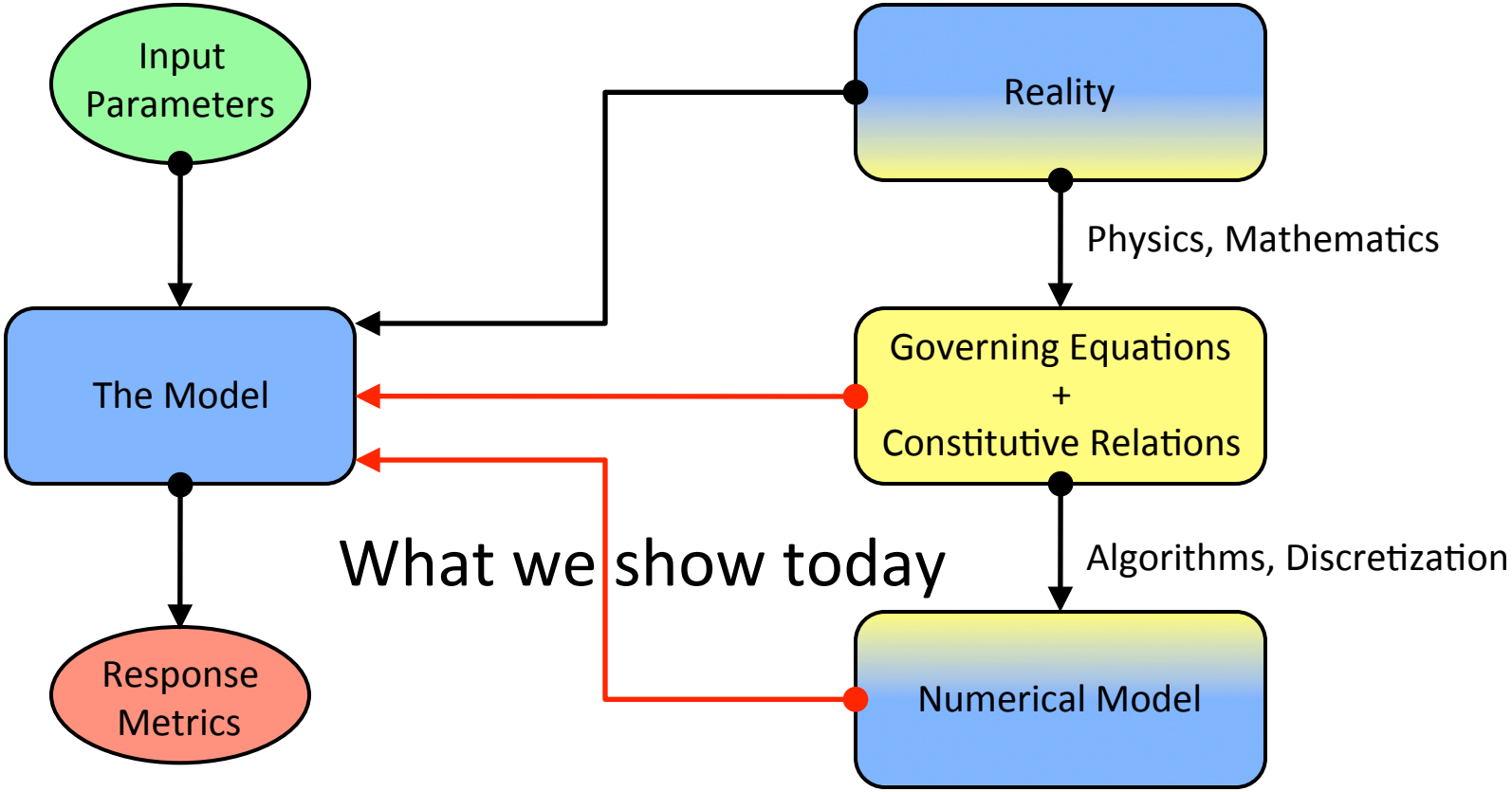
- Hyperbolic Conservation Laws are PDEs that describe the conservation of mass, momentum, and energy.
  - Constitutive relations describing specific materials are also required.
  - This combination is a mathematical model of reality.
- We use algorithms to obtain discrete equations from the mathematical model, and solve the discrete equations using a computer.
  - Such simulations provide approximate numerical solutions to the mathematical model.



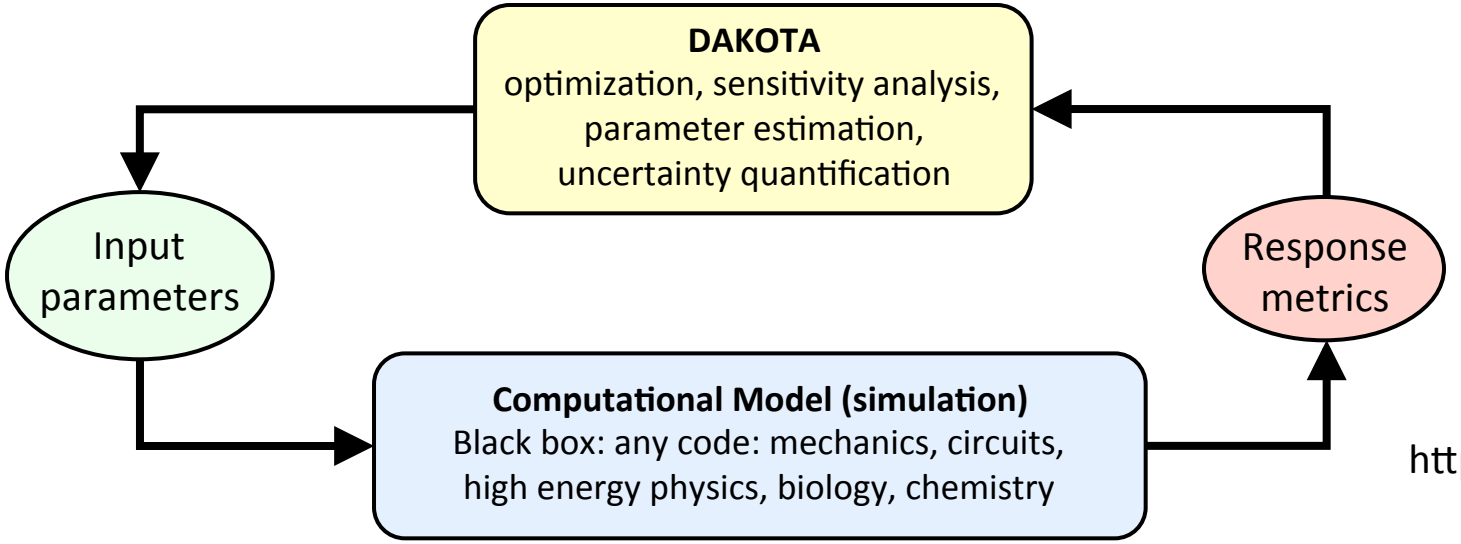
What we get from Sensitivity Analysis of Computer Simulations



What we get from Sensitivity Analysis of Computer Simulations



We conduct sensitivity analyses with DAKOTA.

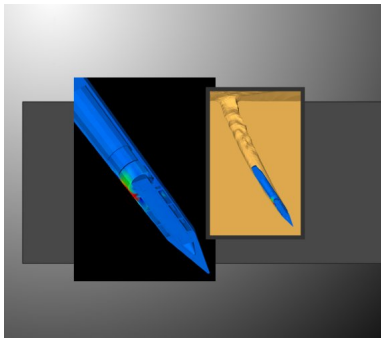


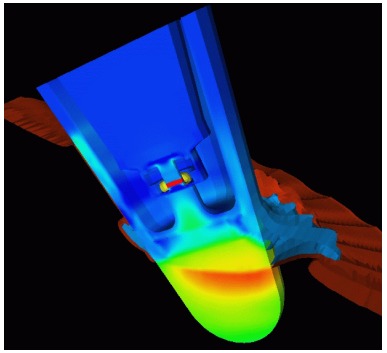
http://dakota.sandia.gov/

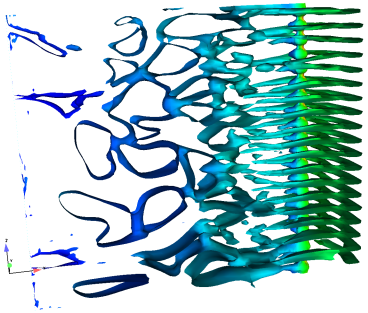
- DAKOTA can automate typical "parameter variation" studies with a generic interface to simulation software and advanced methods.
- UQ methods in DAKOTA include:
  - Sampling (LHS, quasi-MC, classical experimental designs, OAs, VBD)
  - Reliability methods (FORM, SORM, AMV+, etc.)
  - Dempster-Shafer Evidence Theory
  - Stochastic expansion methods: Polynomial chaos, stochastic collocation
  - Epistemic-aleatory nested approaches

We simulate this problem with the ALEGRA multi-physics code.

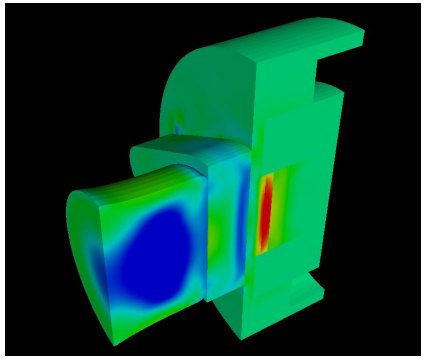
Shock and Multi-physics

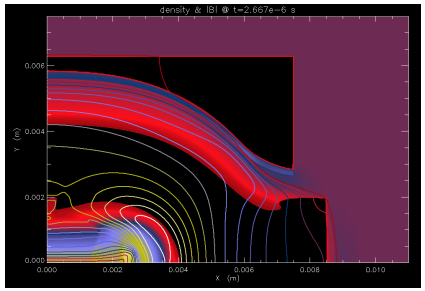


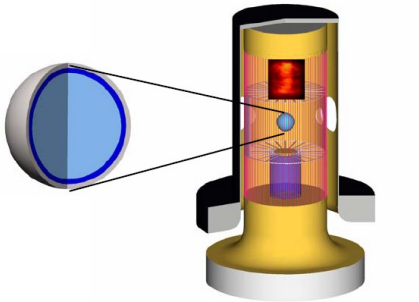




HEDP Theory and ICF Target Design





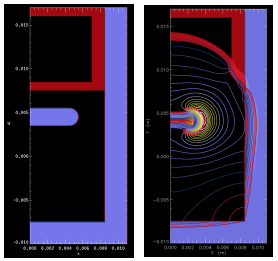


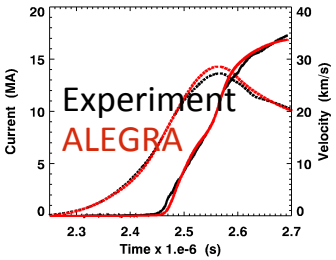
Overview

- The ALEGRA suite of applications models shock and high energy environments for solids, fluids, and plasmas using a multi-material arbitrary Lagrangian-Eulerian (ALE) multi-physics methodology.
- ALEGRA applications run on large, parallel, message-passing architectures in 2-D and 3-D geometries.

ALEGRA Applications

- Shaped Charges & Explosively Formed Penetrators
- Magnetohydrodynamics (MHD)
- Z-pinch, Inertial Confinement Fusion
- Isentropic Compression Experiments/Magnetic Flyers





Isentropic Compression:
Magnetic Flyer Prediction vs. Exp't

Sandia National Laboratories

Hyperbolic conservation laws form the basis of many computational physics investigations.

- Hyperbolic conservation laws are the PDE form of balance laws.
  - For example, the conservation of mass, momentum and energy
- The general form of these equations is given by:

State

$$\frac{\partial \boxed{U}}{\partial t} + \operatorname{div} \boxed{f(U)} = 0 \qquad x \in \Omega \subset \Re^d, \; t \geq 0$$

Flux function

  - The state $U(x,t)\colon \Re^d \times \Re \rightarrow \Re^n$ is the array of conserved quantities
  - There may be additional constraints or source terms on the RHS
  - Appropriate initial and boundary conditions must also be given
- These equations admit complicated solutions with discontinuities.
  - For example, shock waves are governed by equations of this form
- Here, we consider the conservation laws that govern the 1-D equations for compressible fluid flow.

Sandia National Laboratories

An archetypal case for 1-D compressible flow is the experimental "shock tube" configuration.

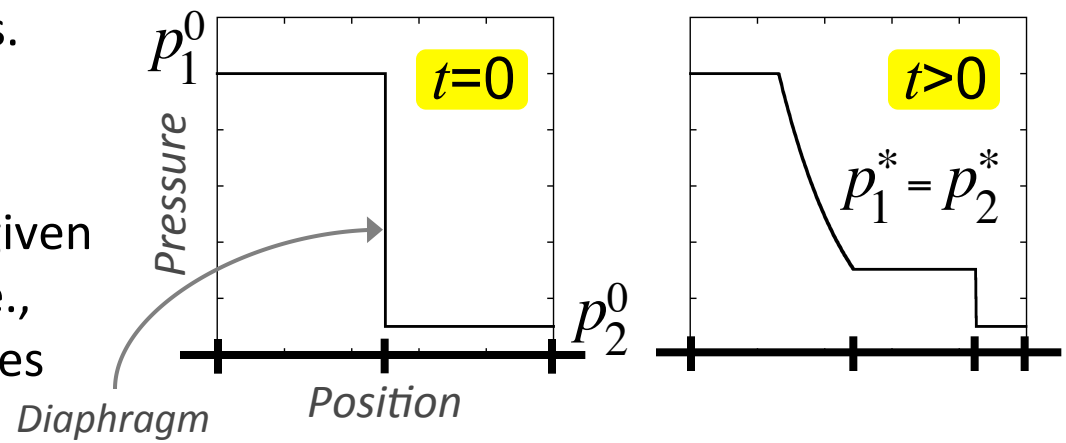
- For 1-D compressible, inviscid, non-heat-conducting flow, the state U and flux f are given by

$$U = [\rho,\ \rho u,\ \rho E]^{\mathsf{T}} \qquad f = [\rho u,\ \rho u^2 + p,\ \rho E u + p u]^{\mathsf{T}}$$

where $E = e + \frac{1}{2}u^2$

Specific internal energy (SIE) *Specific kinetic energy*

- An example is an experimental shock tube for gas dynamics.
  - Conservation laws = PDEs.
  - Constant, uniform initial conditions.
  - For $t > 0$, the solution is given by a set of self-similar (i.e., functions of x/t only) waves



- This is a specific case of the so-called Riemann problem.

What is the "Riemann Problem"?

- 1D gas dynamics equations: $U_t + f_x(U) = 0$

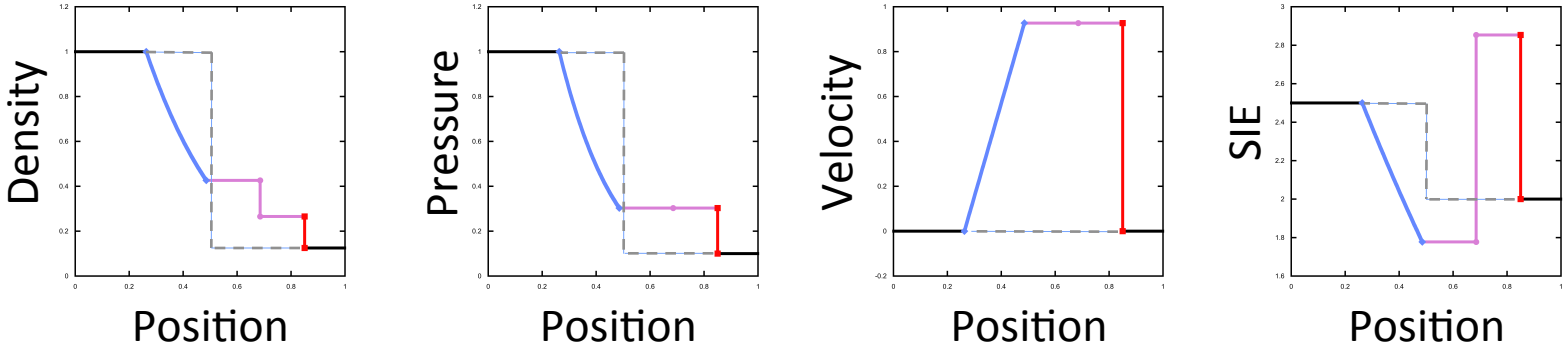
$$U = [\rho,\ \rho u,\ \rho E]^{\mathsf{T}} \qquad f = [\rho u,\ \rho u^2 + p,\ \rho E u + p u]^{\mathsf{T}}$$

with a an ideal gas EOS: $p = p(\rho,\varepsilon) = (\gamma - 1)\,\rho\,\varepsilon$

Constitutive model

G.F. Bernhard Riemann
(1826-1866)
"the greatest mathematician"—*Peter Lax*

- RPs* are the canonical IVP, with two constant initial states, leading to standard solutions:



– Canonical wave structures: Rarefaction, Contact, Shock

*This is a generalization of the concepts introduced by Riemann in "Über die Fortpflanzung ebener Luftwellen von endlicher Schwingungsweite", *Abh. König. Gesell. Wiss. Göttingen, Math.-phys. Klasse*, **8**, pp. 43–65 (1860).

The Riemann problem can have *very different* solutions, depending on the initial conditions.

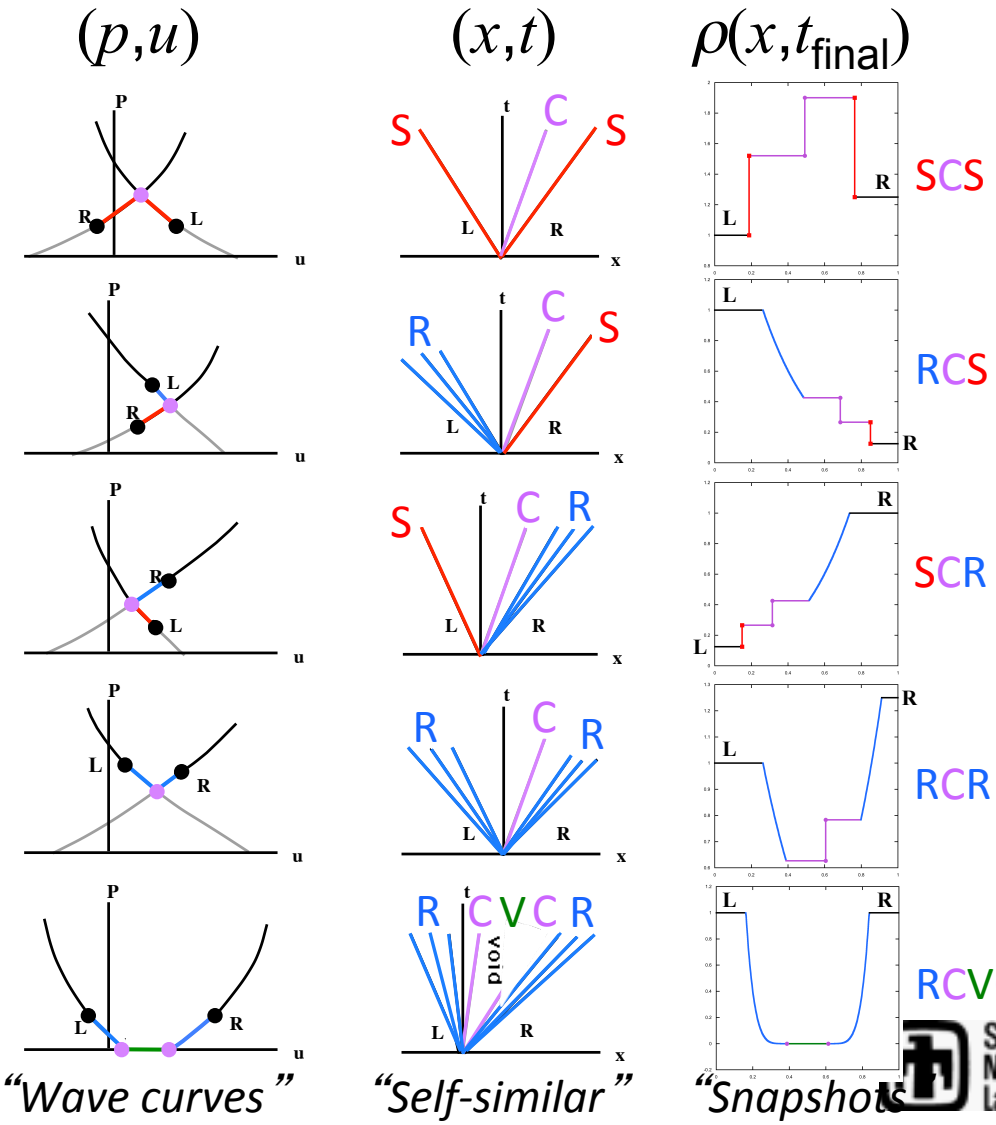
- There are five basic solutions for the 1D gas dynamics equations with an ideal gas EOS*.
  - These depend on the relative pressures and velocities in the ICs
- S = Shock

  C = Contact

  R = Rarefaction

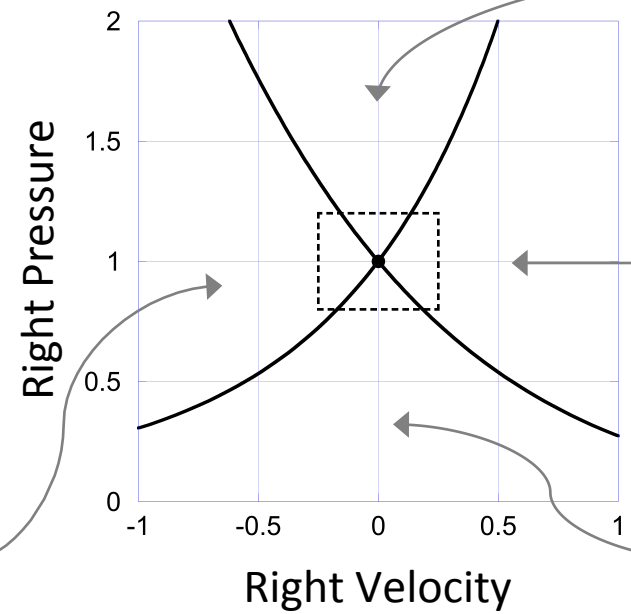
  V = Vacuum/Void

*R. Menikoff, *Applications of Non-Reactive Compressible Fluids*, LANL Report LA-UR-01-273 (2001)

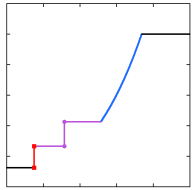


We focus on sensitivity analysis for a single problem, related to the well-known Sod shock tube.

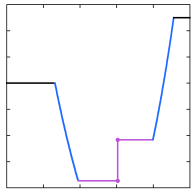
- Initial state: $(\rho,p,u,\gamma) = \begin{cases} (1.0,\ 1.0,\ 0.0, 1.4), & 0 \le x < 0.5 \quad \text{“Left”} \\ (0.125, 1.0, 0.0, 1.4), & 0.5 < x \le 1.0 \quad \text{“Right”} \end{cases}$
- Fix the left state; vary the right state; consider fixed $t_{\text{final}} = 0.2$
- The solution structure varies significantly near the initial point (•).
- Evaluate the sensitivity near that point.



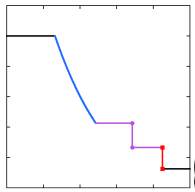
Shock
Contact
Rarefaction

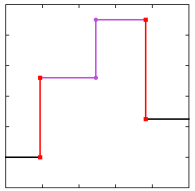


Rarefaction
Contact
Rarefaction



Rarefaction
Contact
Shock





Shock
Contact
Shock

Sod G., : “A Survey of Several Finite Difference Methods for Systems of Nonlinear Hyperbolic Conservation Laws”, *J. Comput. Phys.*, **27**, pp. 1–31 (1978).

Sandia National Laboratories

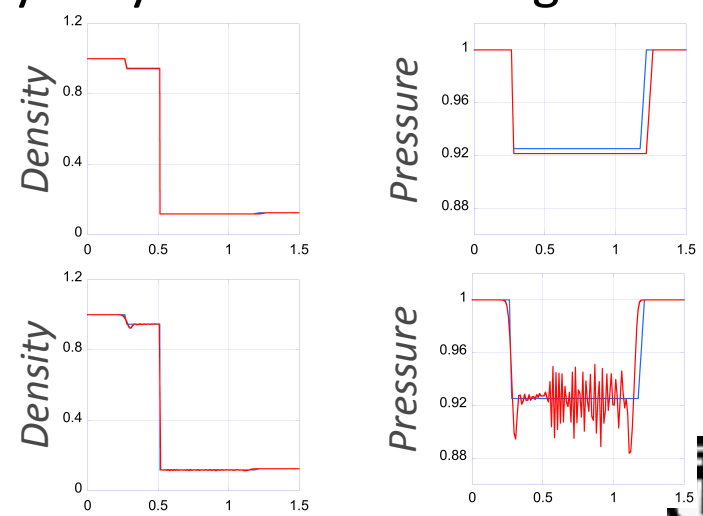
We fix the final time and the left state, but vary both the right state and a numerical parameter.

		Input	*Why?*
Right	X_1	Initial pressure on right	Uncertainty in initial condition
	X_2	Initial velocity on right	Uncertainty in initial condition
	X_3	Polytropic index γ on right	Uncertainty in material model
	X_4	CFL parameter: $c_s \, \Delta t / \Delta x$	Numerical parameter

- From the self-similar nature of the solution, only one state need be varied, not both: hence, we vary only values on the right.

- Higher pressure, higher γ $\rightarrow$ higher sound speeds and faster wave propagation

- $0 < \text{CFL} < 1 \rightarrow$ stable
  $\text{CFL} > 1 \rightarrow$ unstable

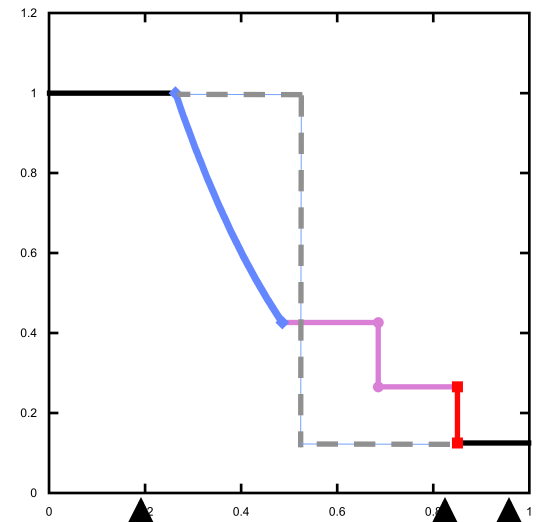


Nominal
High γ

Nominal
High CFL

Most of our responses are from probes

- A probe measures some quantity at some location
  - We measured at one location on the left and two on the right of the initial interface location
  - We record the value at the end of the simulations, t=0.2



X=0.35
Y5, Y6, Y7

X=1.16
Y2, Y3, Y4

X=1.4
Y1

We consider specific characteristics of the solution as the output variables.

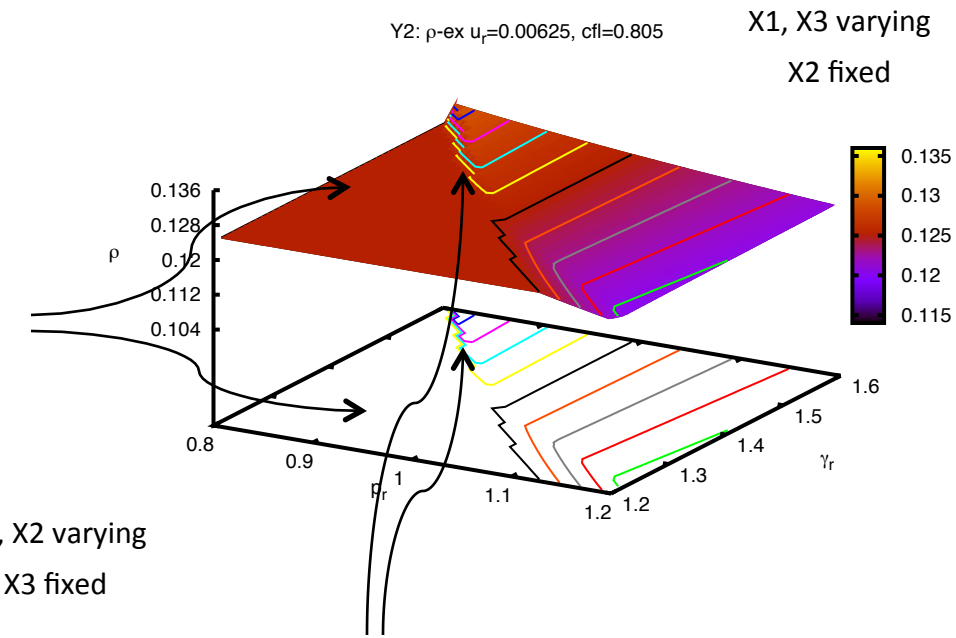
		Output	*Why?*
Right	Y_1	Specific internal energy, x = 1.4	Coupled physics
	Y_2	Mass density, x = 1.16	Wave speed
	Y_3	Kinetic energy, x = 1.16	Physics diagnostic
	Y_4	Time of 1st $\Delta\rho$, x = 1.16	Experimental diagnostic
Left	Y_5	Mass density, x = 0.35	Wave speed
	Y_6	Kinetic energy, x = 0.35	Physics diagnostic
	Y_7	Time of 1st $\Delta\rho$, x = 0.35	Experimental diagnostic
	Y_8	CPU time	Computational diagnostic

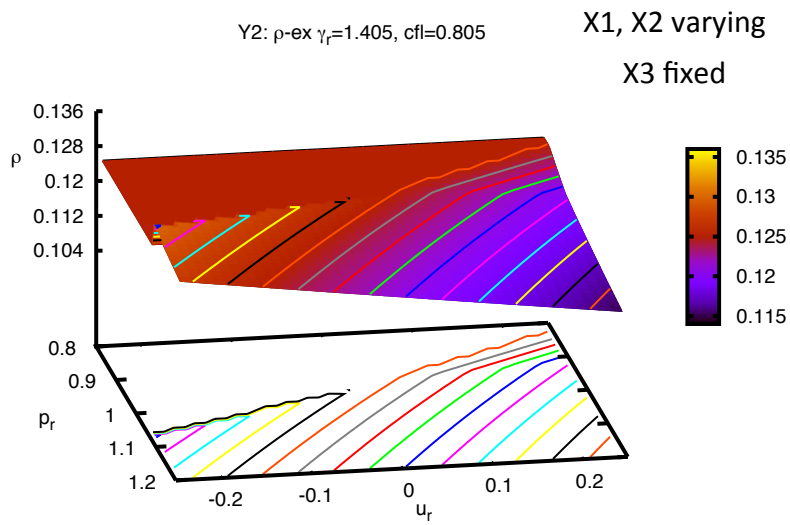
- Shock-Physics analysts think of the problem in these terms

Sandia National Laboratories

Y2: ρ at x = 1.16
Output surface slices for the Exact Model

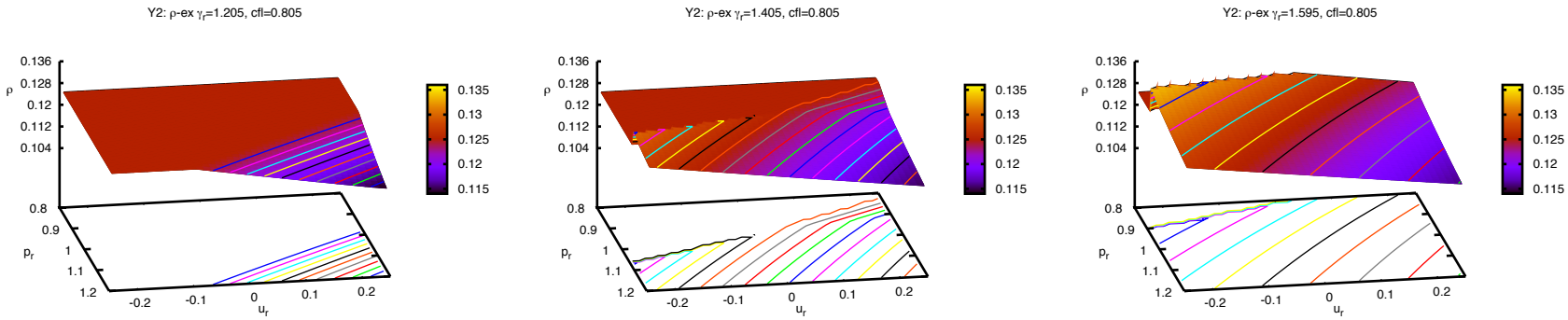
Flat plateaus indicate no waves have reached this location





Sharp jumps indicate shocks

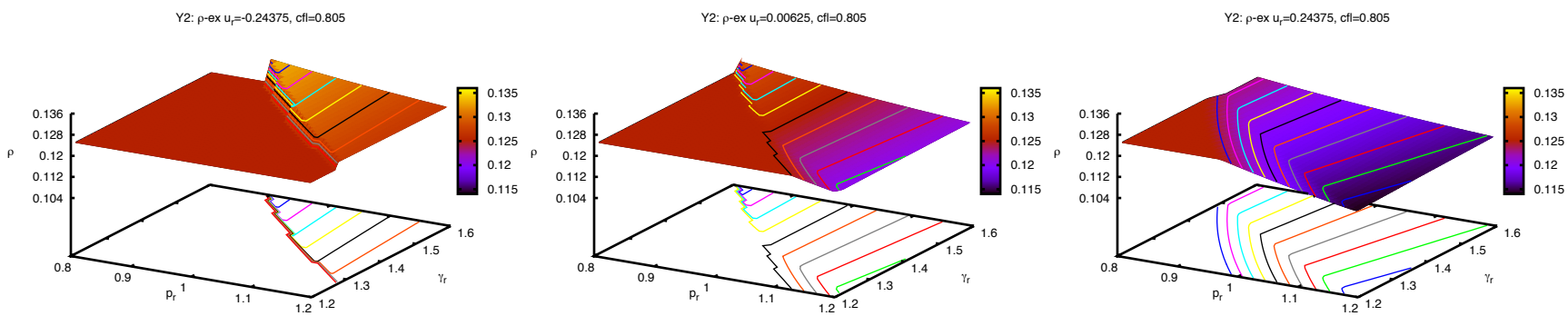
Y2: ρ at x = 1.16
Output surface slices for the Exact Model



X3 increasing

X1, X2 varying

X3 fixed

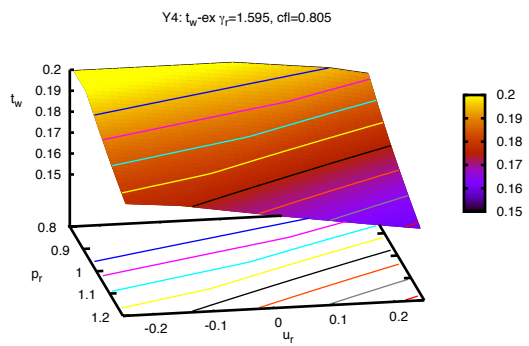


X2 increasing

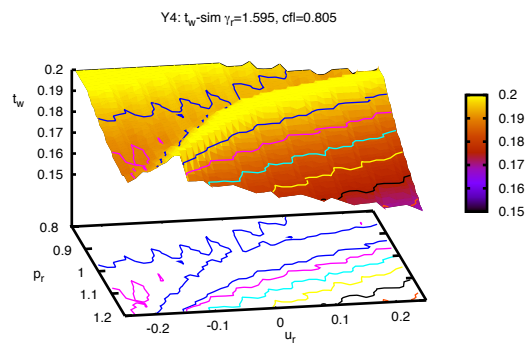
X1, X3 varying

X2 fixed

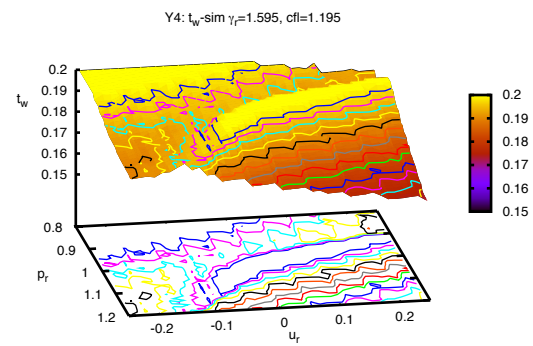
Y4: t_w at x = 1.16
Output surface slices for the Simulation Model



Exact Model



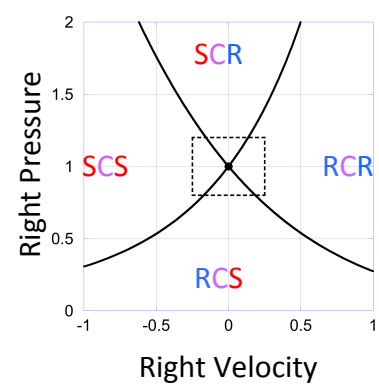
Simulation Model X4 = 0.805



Simulation Model X4 = 1.195

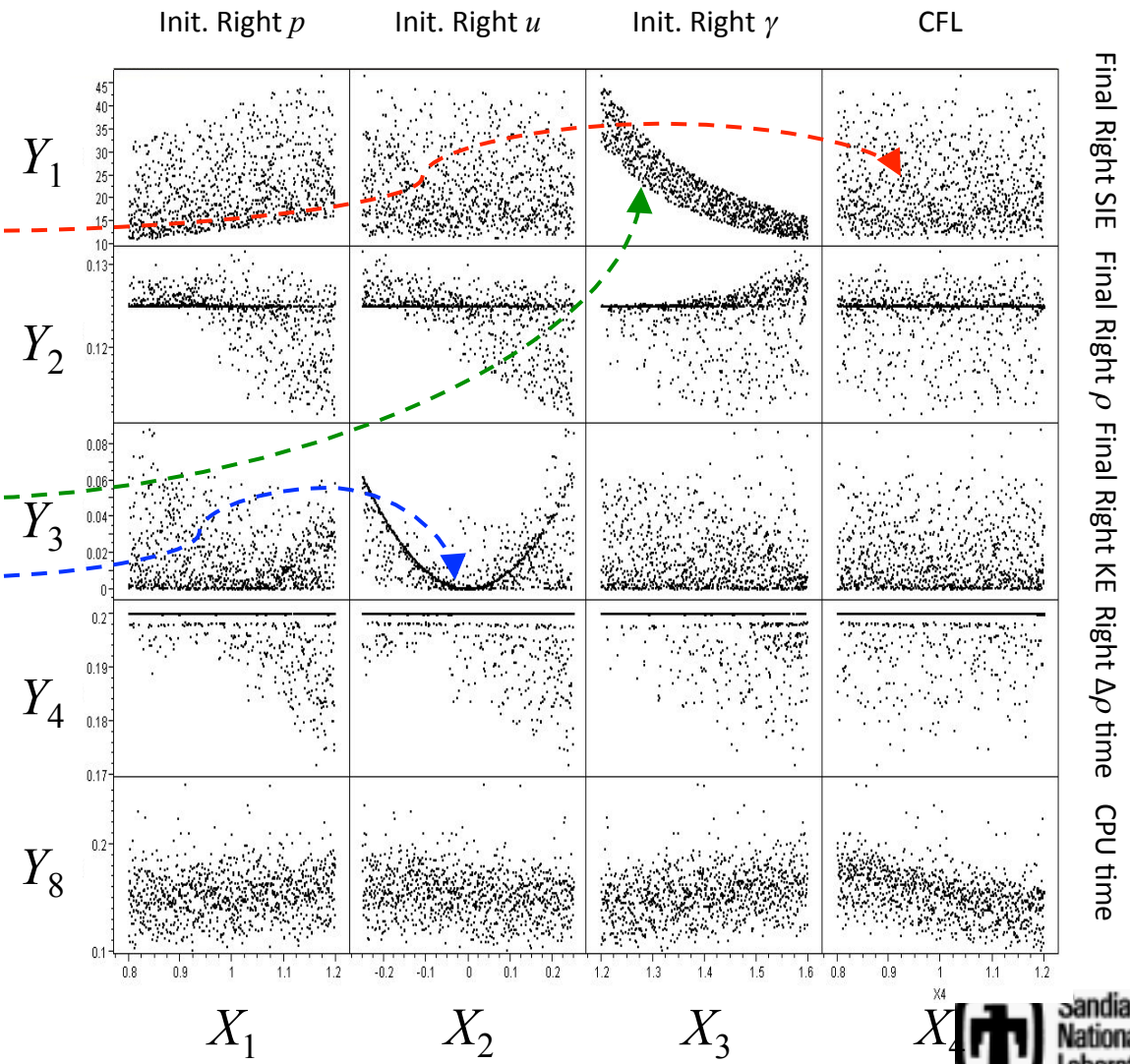
X1, X2 varying X3, X4 fixed

- In a few cases, Simulation Model response surfaces show a different topology than the Exact Model



Scatterplots of model outputs give some insights into the distributions.

- Some outputs appear insensitive:
  - Y_1 variation with X_4
- Some trends in the data are clear:
  - Y_1 variation with X_3
  - Y_3 variation with X_2
- VBD sensitivity indices quantify this behavior…

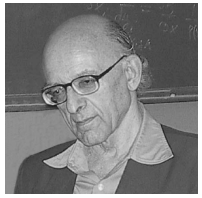


Correlation and Variance-Based Decomposition (VBD) are global sensitivity characterizations of uncertainty in model outputs Y.

- Goal: to assess inputs over a hypercube of interest.
- Correlation analysis identifies the strength and direction of a *linear* relationship between input and output.
- VBD identifies the fraction of the variance in the output that can be attributed to an individual variable alone or with interaction effects.
  - Main effect sensitivity S_i is the fraction of the uncertainty in Y that can be attributed to input x_i *alone*
  - Total effect index T_i is the fraction of the uncertainty in Y that can be attributed to x_i *and its interactions with other variables*

$$S_i = \frac{\mathrm{Var}_{x_i}\left[E(Y|x_i)\right]}{\mathrm{Var}(Y)}$$

$$T_i = \frac{E\left[\mathrm{Var}(Y|x_{-i})\right]}{\mathrm{Var}(Y)}$$



I.M. Sobol' developed these ideas

  - Calculation of S_i and T_i requires the evaluation of m-dimensional integrals, approximated by Monte-Carlo sampling.

$$x_{-i} \equiv (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_m)$$

  - *Computationally intensive*, as replicated sets of samples are evaluated: N samples and D inputs→ evaluation of $N \times (D+2)$ samples.

How sensitivity indices are calculated

$$S_i = \frac{Var(E(Y \mid X_i))}{Var(Y)}$$

- Full Factorial:
  - Take n values of each input variable X_i; the number of samples are a full tensor product of n samples in each input variable, $N = n^d$
  - For each particular value of X_i, calculate the average over the other X_j variables.
  - Calculate the variance of this expectation (variance over n values)

$$E(Y \mid X_i = x_{ik})$$

$$Var(E(Y \mid X_i))$$

- Approximation in *Sensitivity Analysis in Practice* (Satelli et al. 2004):
  - Calculate two independent sample matrices, A and B, with d (number of inputs) columns and n rows . C_i is constructed by taking the ith column of A and substituting it into B.
  - Y_A, Y_B, and Y_{Ci} are the vectors of responses from evaluating the simulator at the sample values in A, B, or C_i.
  - Total samples is (2+d)*n
  - Requires that n is of order thousands for accuracy

$$f = \frac{Y_A \bullet Y_B}{N}$$

$$estimated\,\mathrm{var}(Y) = \left(\frac{1}{N-1} Y_A \bullet Y_A\right) - f^2$$

$$S_i = \frac{\left(\frac{1}{N-1} Y_A \bullet Y_{c_i}\right) - f^2}{estimated\,\mathrm{var}(Y)}$$

Sandia National Laboratories

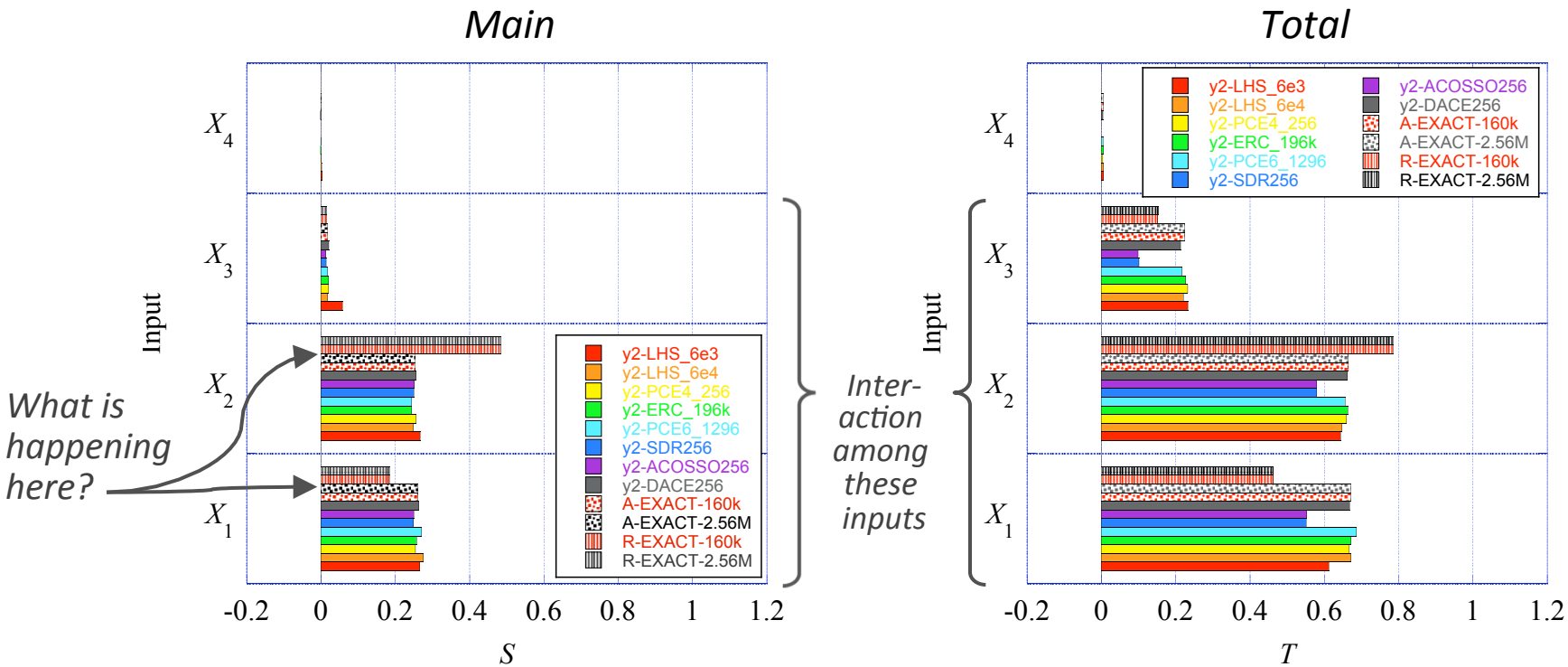
Comments on Computing Sensitivity Indices

- The multidimensional integrals are usually computed approximately – estimated from limited information. In practice, this isn't so bad because the indices are used primarily for ranking inputs rather than precise comparisons.
- Nonetheless, it is still important to know how the accurate the estimates are as a function of, e.g., sample size, and that they converge to the correct value.
- Exact solution of the sensitivity analysis problem: full factorial "sampling" of the uncertain inputs can be viewed as numerical quadrature for uniformly distributed inputs.

Sandia National Laboratories

The sensitivity indices for Y_2 have some interesting features.

- For Y_2 (final right ρ), main and total indices have different values, indicating those inputs interact with others.

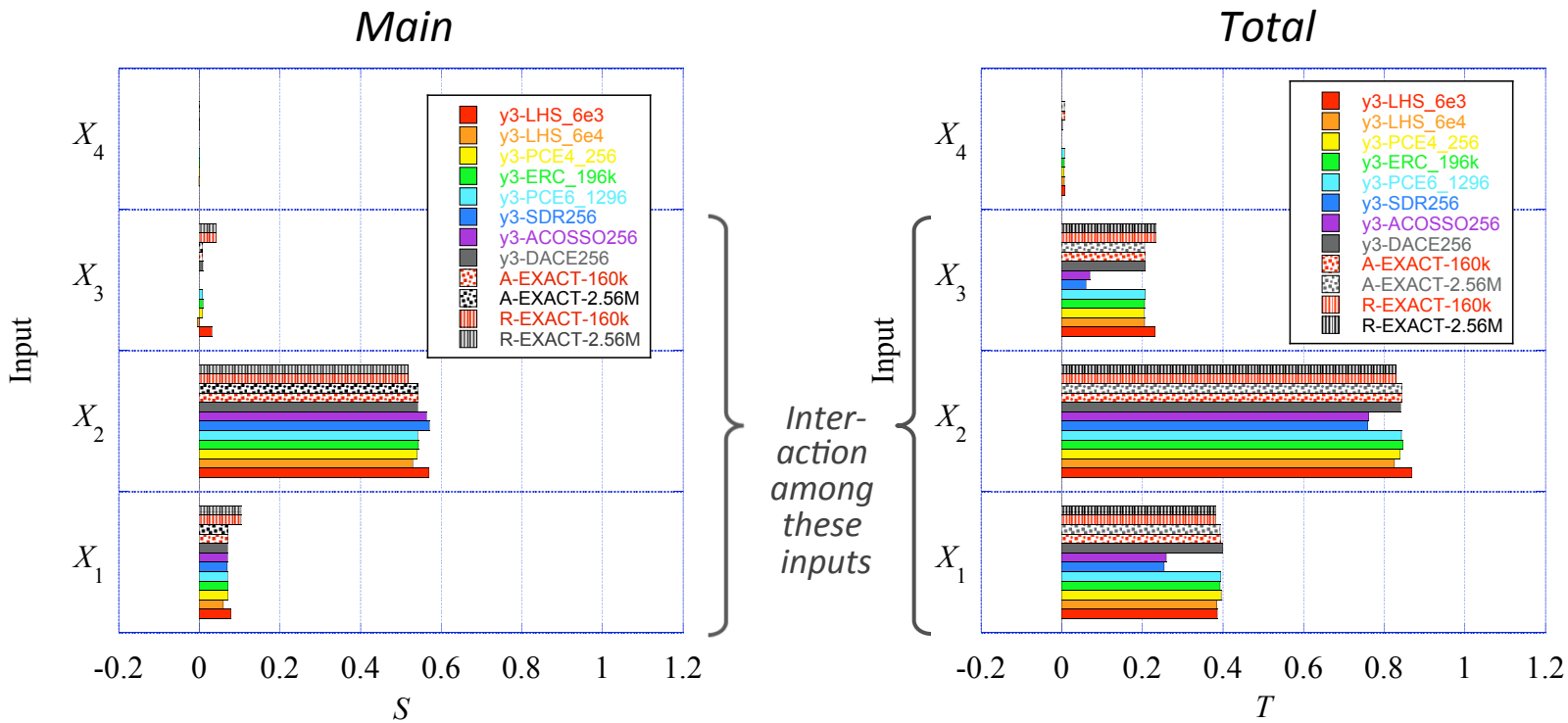


LHS 6000 PCE4 256 PCE6 1296 ACOSSO 256 A-EXACT 160k R-EXACT 160k

LHS 60000 JRC 196k SDP 256 DACE 256 A-EXACT-2.56M R-EXACT-2.56M

The sensitivity indices for Y_3 perform similarly for all approaches.

- As anticipated, Y_3 (final right KE) depends strongly on X_2 (u_R).
  - Sensitivity on X_3 (γ_R) is less than heuristically expected.



LHS 6000 PCE4 256 PCE6 1296 ACOSSO 256 A-EXACT 160k R-EXACT 160k

LHS 60000 JRC 196k SDP 256 DACE 256 A-EXACT-2.56M R-EXACT-2.56M

What are we looking at?

- Two different models -> two different sensitivity analysis problems
  1. Model 1: exact solution to the Riemann problem
  2. Model 2: approximate numerical solution to the Riemann problem (Alegra simulations)

- Solutions for each sensitivity analysis problem are essentially exact.

Sandia National Laboratories

What else could we have done?

- Treat simulation model discretization parameters as uncertain inputs and compute sensitivity indices. This is a common practice; *more troubling is that it is often viewed as sufficient code verification*.

- Consider the sensitivity of the differences between the exact and simulation models
- Consider error of the code verification problem as sensitivity analysis outputs.

Sandia National Laboratories

Sensitivity Indices for y2 Differences

Main Effects	Y2-a	Y2-r	Y2-diff
S_4	0.0017	0.0	0.0011
S_3	0.0175	0.0218	0.0544
S_2	0.2487	0.4858	0.0688
S_1	0.2757	0.1865	0.0602

Total Effects	Y2-a	Y2-r	Y2-diff
T_4	0.0050	0.0	0.0173
T_3	0.2214	0.1529	0.7621
T_2	0.6477	0.7704	0.6495
T_1	0.6715	0.4523	0.8644

- a: alegra
- r: Riemann solver
- diff: error

- diff is not the difference between Y2-a and Y2-r; it is the sensitivity index of Y2-a – Y2-r

- Results from LHS-60k

Sandia National Laboratories

Sensitivity Indices for y3 Difference

Main Effects	Y3-a	Y3-r	Y3-diff
S_4	-0.0010	0.0	0.0023
S_3	-0.0043	0.0440	0.0737
S_2	0.5296	0.5046	0.2811
S_1	0.0598	0.0927	0.0048

Total Effects	Y3-a	Y3-r	Y3-diff
T_4	0.0063	0.0	0.0183
T_3	0.2063	0.2354	0.6471
T_2	0.8279	0.8097	0.7988
T_1	0.3858	0.3630	0.6018

- a: alegra
- r: Riemann solver
- diff: error

- diff is not the difference between Y3-a and Y3-r; it is the sensitivity index of Y3-a – Y3-r

- Results from LHS-60k

It is not obvious how to interpret the sensitivity of the differences between the exact and simulation models.

Sandia National Laboratories

Sensitivity Indices for L_1 Error Norms

Main Effects	dens	press	SIE	velocity
S_4	0.1938	0.1908	0.1707	0.2249
S_3	0.3615	0.5744	0.5414	0.5485
S_2	0.0905	0.0191	0.0462	0.0259
S_1	0.1948	0.0626	0.0858	0.0473

Total Effects	dens	press	SIE	velocity
T_4	0.2824	0.2893	0.2625	0.3301
T_3	0.4534	0.6853	0.6528	0.6611
T_2	0.1577	0.0539	0.0869	0.0607
T_1	0.2644	0.1345	0.1703	0.1088

Results from LHS-60k

Sandia National Laboratories

What else can we say given two alternative models?

That is, we have a simulation model and an exact model.

We know how to use the exact model for code verification on a point by point (in input parameter space) basis.

- How can we use UQ techniques to perform code verification over a broad input parameter space?
- Can such techniques help us filter “noisy” convergence rates? How much deviation from design convergence rates is acceptable?

Sandia National Laboratories

Cook-Cabot Convergence Rates

L1 density error convergence rates, Cook-Cabot problem

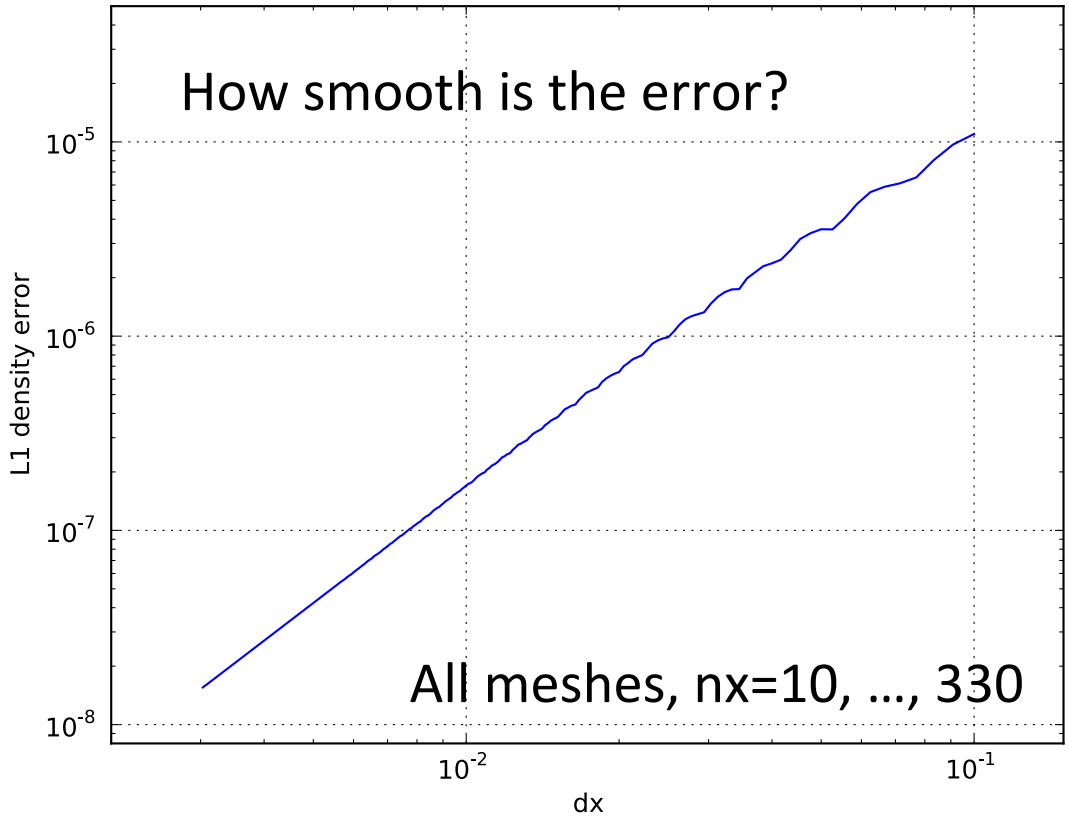
NX	18		36		72		144		288
Order		1.675		1.957		2.003		1.999	
NX	19		38		76		152		304
Order		1.626		1.982		1.984		2.004	
NX	20		40		80		160		320
Order		1.846		1.888		2.012		2.001	

Why do the convergence rates vary so much for the three similar sequences?

Cook, A. W. and Cabot, W. H., : "A High-Wavenumber viscosity for High-Resolution Numerical Methods", *J. Comput. Phys.,* **195**, pp. 594–601 (2004).

Cook-Cabot Density Error Norms

L1 density error, Cook-Cabot problem



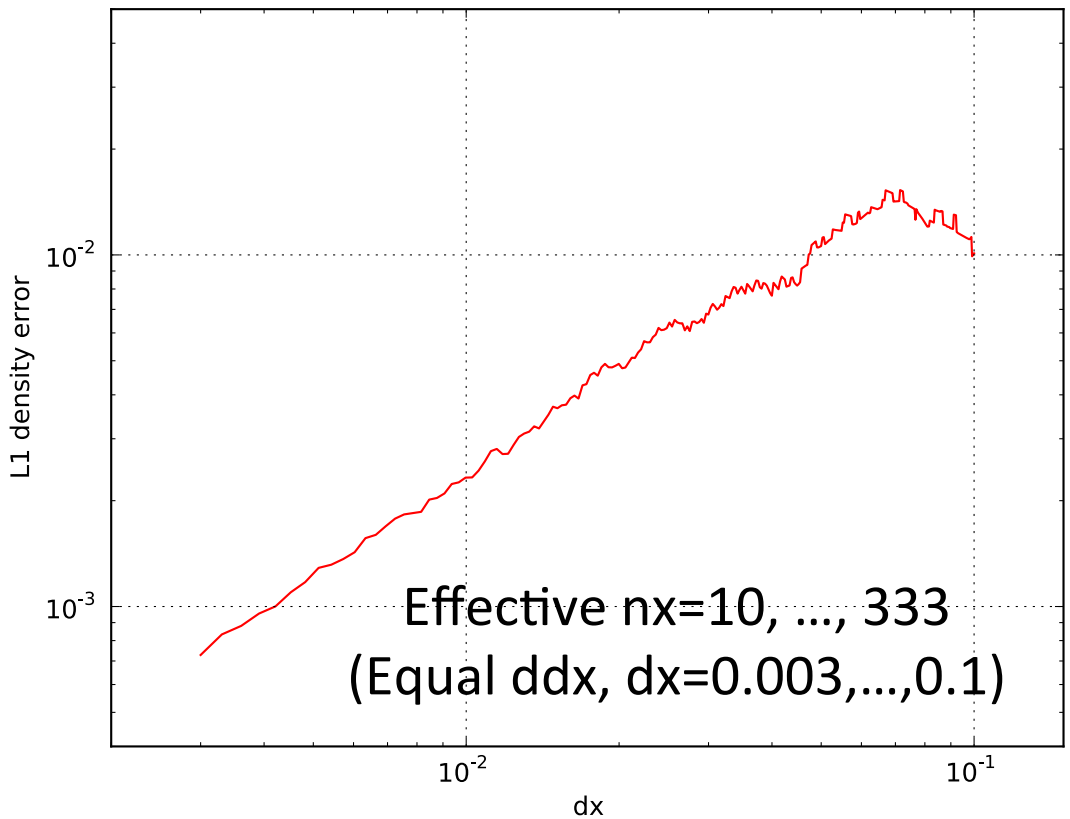
Sod Convergence Rates

L1 density error convergence rates, Sod problem

NX	18		36		72		144		288
Order		0.972		0.984		0.930		1.025	
NX	19		38		76		152		304
Order		0.846		0.985		1.034		0.962	
NX	20		40		80		160		320
Order		0.803		1.062		0.973		0.975	

Why do the convergence rates vary so much for the three similar sequences?

Sod G., : "A Survey of Several Finite Difference Methods for Systems of Nonlinear Hyperbolic Conservation Laws", *J. Comput. Phys.*, **27**, pp. 1–31 (1978).

Sod Density Error Norms

L1 density error, Sod problem