Approved for Unlimited Release
SAND2010-XXXXC

SAND2010-6795C

EUROPEAN COMMISSION

A Sensitivity Analysis of the Gas Dynamics Equations

James R. Kamm*, V. Gregory Weirs*, Laura P. Swiler*, Marco Ratto§, Stefano Tarantola§, Brian M. Adams*, William J. Rider*, Michael S. Eldred*

* Sandia National Laboratories, Albuquerque, NM USA
§ Joint Research Center, European Commission, Ispra (VA), Italy

XII International Workshop on Supercomputing and Mathematical Modeling

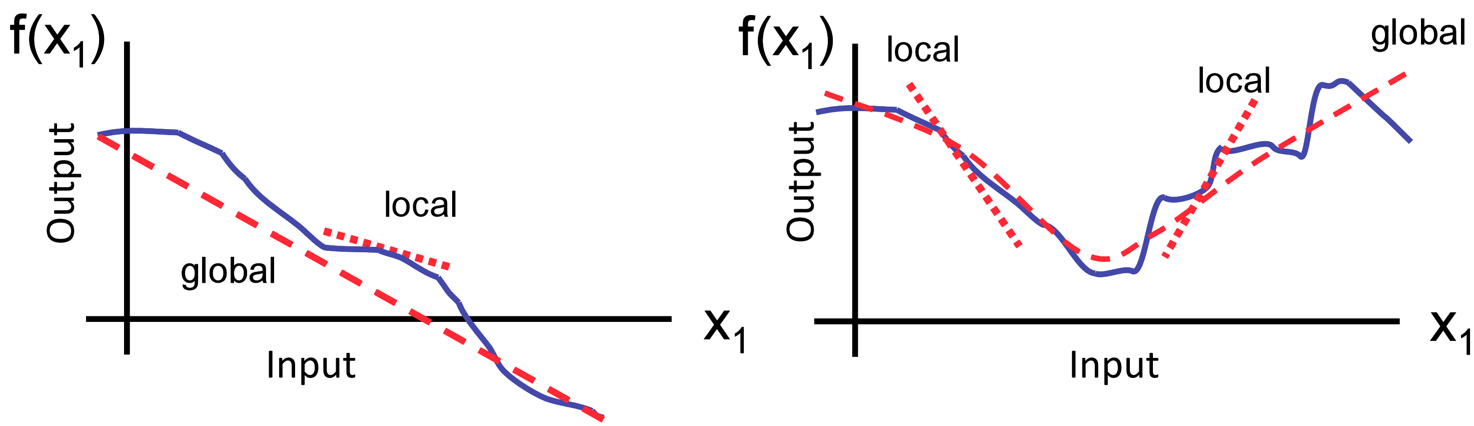
11–15 October 2010, Sarov, Russia

Sandia National Laboratories is a multi-program laboratory operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin company, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

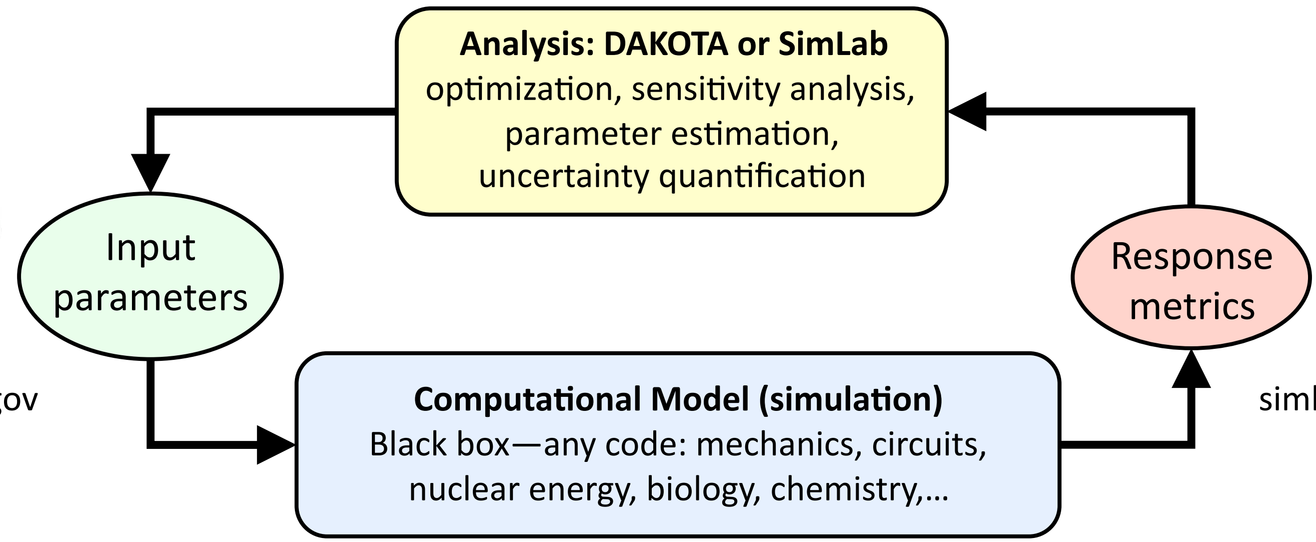
What is Sensitivity Analysis?

- Sensitivity Analysis (SA) is a way to order the input variables to a model according to their relative importance to the model's output.
- The results of SA can be used to inform us about:
  - Optimization
  - Uncertainty Quantification
  - Which inputs to gather more data on
  - How to better control an experiment
- <u>Local</u> SA: local linear or under-resolved behavior can be misleading.
- <u>Global</u> SA: can be computationally expensive (use meta-modeling).



Sandia National Laboratories

We conduct sensitivity analyses with DAKOTA and SimLab.



dakota.sandia.gov

simlab.jrc.ec.europa.eu

- DAKOTA has a generic interface to simulation software, contains advanced methods, and can automate "parameter variation" studies, including:
  - Sampling (LHS, quasi-MC, classical experimental designs)
  - Dempster-Shafer evidence theory
  - Stochastic expansion methods: Polynomial chaos, stochastic collocation
  - Nested approaches for quantifying epistemic and aleatory uncertainties
- SimLab is a development framework designed for Monte Carlo-based uncertainty and sensitivity analysis.

Sandia National Laboratories

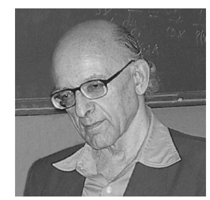
Correlation and Variance-Based Decomposition (VBD) characterize the global sensitivity of model outputs Y to model inputs X.

- Goal: to assess outputs for a specified range of inputs.
- 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 is due to input X_i *alone*

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

  - Total effect index T_i is the fraction of the uncertainty in Y that is due to X_i *and its interactions with other variables*

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



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.

$$X_{-i} \equiv (X_1, \ldots, X_{i\text{-}1}, X_{i+1}, \ldots, X_m)$$

  - The evaluation of these quantities is *computationally intensive*, as replicated sets of samples are evaluated.

Sandia National Laboratories

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: total # = N^D
  - For each particular value of X_i, calculate the average over the other X_j variables. $E(Y \mid X_i = x_{ik})$
  - Calculate the variance of this expectation (variance over N values) $Var(E(Y \mid X_i))$
- 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 i^{th} 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: $N \times (2+D)$.
  - Typically, $N = O(10^3)$ for accuracy

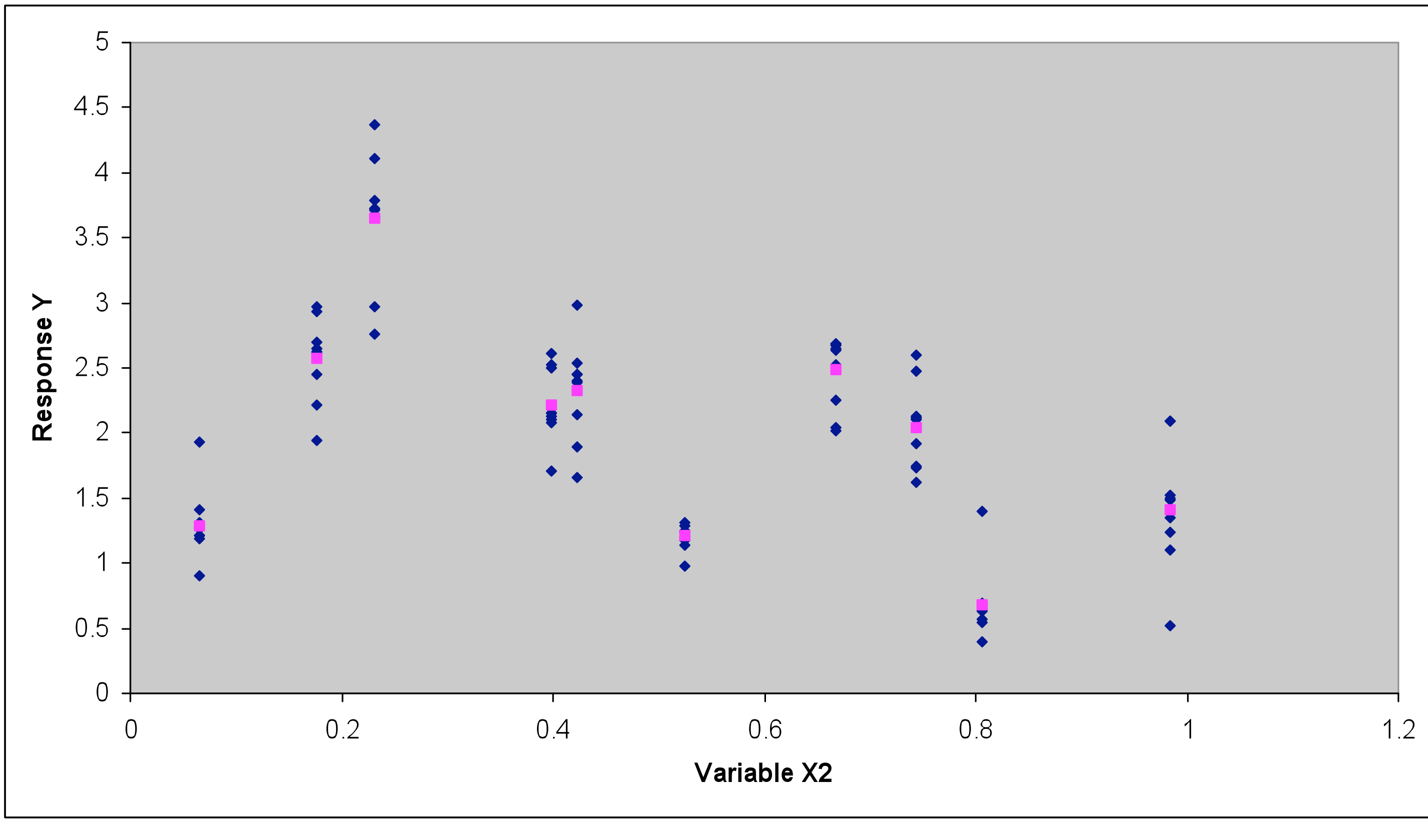
$$f = \frac{Y_A \bullet Y_B}{N}$$

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

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

Variance-Based Decomposition: a notional example.

- Main effects indices S_j identify the fraction of uncertainty in the output attributed to X_j alone
- Total effects indices T_j corresponds to the fraction of the uncertainty attributed to X_j and its interactions with other variables



Sandia National Laboratories

Meta-models provide an alternative to sampling-based VBD.

- Build the meta-model using some of the data.
  - This is reasonable for moderately high dimensional data
- Estimate the VBD sensitivity indices using the meta-model.
  - This can be done (i) with the same data used to construct the meta-model or (ii) with data that was *not* used to construct the meta-model
- Meta-models can be used to generate confidence intervals of the computed indices.
  - These confidence intervals give a measure of the "variability" or "uncertainty" in the computed indices.
- There are several different ways to construct the meta-models.
  - "Regression surfaces" (regression and smoothing)
  - Stochastic expansions (polynomial chaos, stochastic collocation)

Sandia National Laboratories

Regression surface models are alternatives to sampling-based approaches.

- SDP = State-Dependent Parameter Regression
  - SDP modeling* is a class of non-parametric smoothing, first suggested by Young§, that is similar to smoothing splines and kernel regression approaches but is performed using recursive (non-numerical) Kalman filter and associated fixed interval smoothing.
  - SDP is good for additive models and can adapt to local discontinuities, strong non-linearity, and heteroskedasticity in the response.
- ACOSSO = Adaptive COmponent Selection and Smoothing Operator
  - ACOSSO† is a multivariate smoothing-spline approach (COSSO‡) that is augmented by a weighted (w_j), scaled (λ) penalty function:

$$\hat{f} = \min_{f \in \mathcal{F}} \left\{ \frac{1}{N} \sum_{i=1}^{N} \left(Y_i - f(x_i) \right)^2 + \lambda \sum_{d=1}^{D} w_d \left\| P^d f \right\| \right\}$$

D = # inputs

  - ACOSSO is thought to perform best for a reasonably smooth underlying response.
- DACE = Design and Analysis of Computer Experiments
  - Gaussian Process emulator for the output responses.

§ Young, P. C. "The identification and estimation of nonlinear stochastic systems," in *Nonlinear Dynamics and Statistics*, A. I. Mees et al., eds., Birkhauser, Boston (2001).

* Ratto, M., Pagano, A., Young, P. C., "State dependent parameter meta-modelling and sensitivity analysis," *Comput. Phys. Comm.*, **177**, pp. 863–876 (2007).

† Storlie, C.B., Bondell, H.D., Reich, B.J., Zhang, H.H., "Surface estimation, variable selection, and the nonparametric oracle property," *Stat. Sinica,* to appear (2010).

‡ Y. Lin,Y., and H. Zhang, H., "Component selection and smoothing in smoothing spline analysis of variance models," Ann. Stat., **34**, pp. 2272–2297 (2006).

Stochastic Expansion Methods provide another alternative to sampling-based VBD.

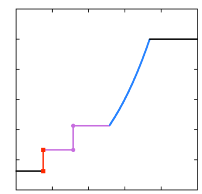
- Stochastic expansion methods — Polynomial Chaos Expansion (PCE) or Stochastic Collocation (SC) — produce functional representations of stochastic variability.
- Sudret* (i) demonstrated that the sensitivity indices are explicit functions of the stochastic expansion, and (ii) derived the PCE case.
  - Once the PCE is obtained, sensitivity indices are calculated *explicitly*, i.e., without additional sampling
- Tang§ derived the sensitivity indices as analytic functions of SC.
- Both of these techniques have been implemented in DAKOTA.
- This approach can be *very efficient*, since the calculation of sensitivity indices does *not* require additional function evaluations.

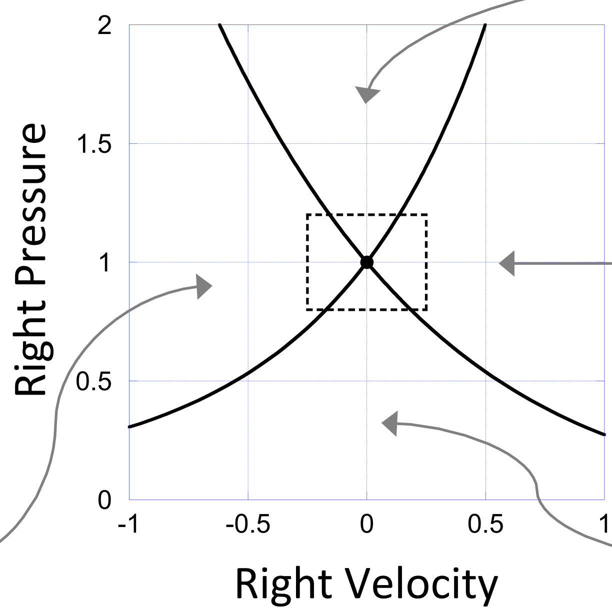
* Sudret, B., "Global Sensitivity analysis using polynomial chaos expansion," *Rel. Engr. & Syst. Safety*, **93**, pp. 964–979 (2008).

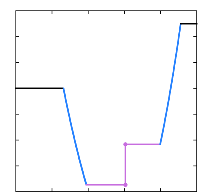
§ Tang, G., Iaccarino, G., Eldred, M.S., "Global Sensitivity Analysis for Stochastic Collocation Expansion," paper AIAA-2010-2922 in *Proceedings of the 12th AIAA Non-Deterministic Approaches Conference*, Orlando, FL, 12–15 April 2010.

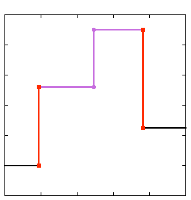
We consider Sensitivity Analysis of a shock tube problem.

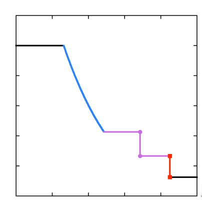
- Initial state: $(\rho, p, u, \gamma) = \begin{cases} (1.0, \ 1.0, 0.0, 1.4), & 0 \le x < 0.5 \quad \text{“Left”} \\ (0.125, 1.0, 0.0, 1.4), & 0.5 < x \le 1.0 \quad \text{“Right”} \end{cases}$
- Fix the left state; vary the right state; consider fixed $t_{\text{final}} = 0.2$
- We can evaluate the exact solution to this problem.
- Examine the sensitivity near the point (•), where the solution varies...











Sandia National Laboratories

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

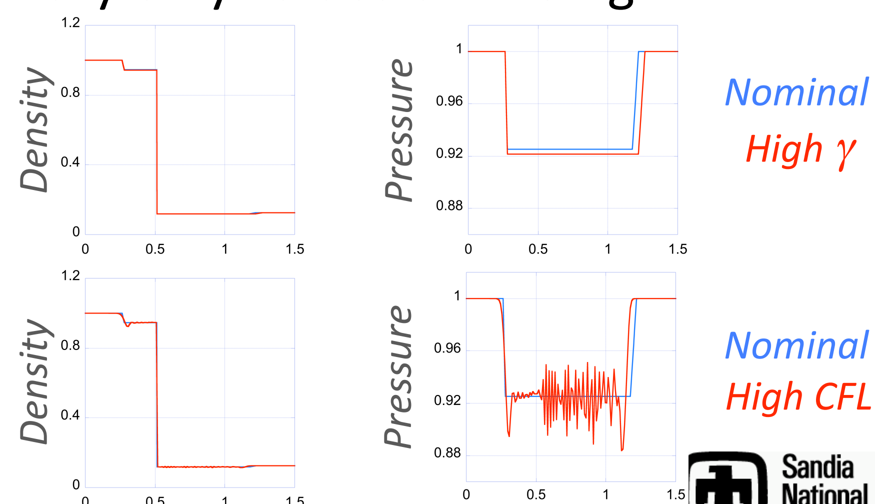
		Input	Why?
Right	X_1	Initial pressure on right	Uncertainty in initial condition
Right	X_2	Initial velocity on right	Uncertainty in initial condition
Right	X_3	Polytropic index γ on right	Uncertainty in material model
	X_4	CFL parameter: $c_s \, \Delta t / \Delta x$	Numerical parameter

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

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

- 0 < CFL < 1 $\rightarrow$ stable

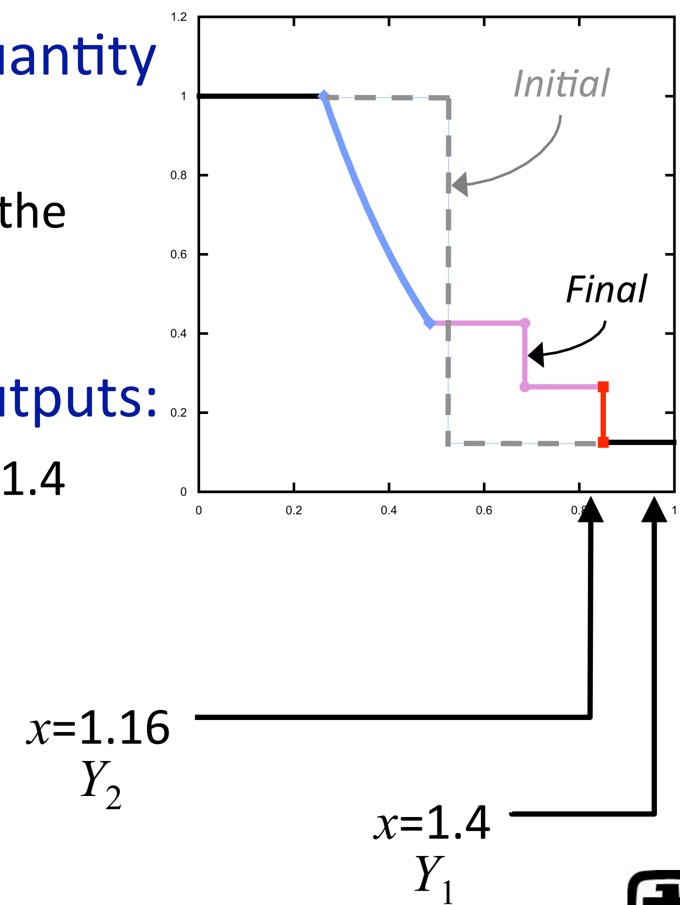
  CFL > 1 $\rightarrow$ unstable



Sandia National Laboratories

The outputs are values at fixed locations.

- These outputs correspond to a experimental diagnostics.
- These outputs measure some quantity at a specific location.
  - We record the value at the end of the simulations, t=0.2
- In this study, we examine two outputs:
  - Y_1 = specific internal energy at x = 1.4
  - Y_2 = density at x = 1.16

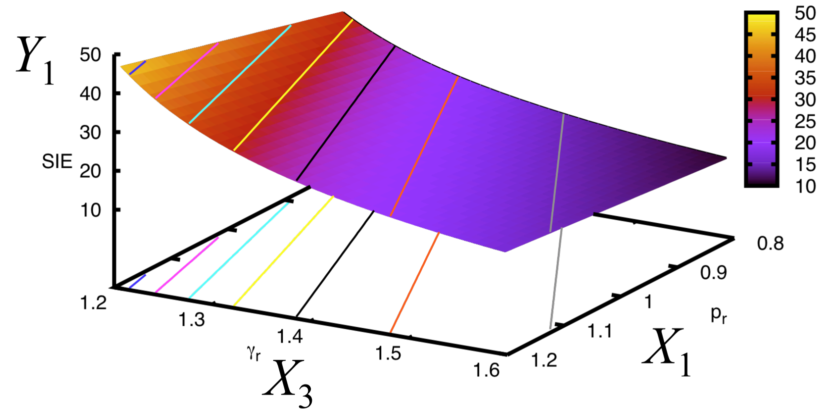


We can evaluate the exact solution for all outputs, e.g., Y_1, which is the SIE at $x = 1.4$.

- Y_1 is a simple output that we use as a test.
- No waves reach this location, so the SIE does not change from at initial value.
  - This value is a function of X_1 and X_3 only.
  - Sensitivity indices should show this dependence.

X_1, X_3 varying

X_2 fixed

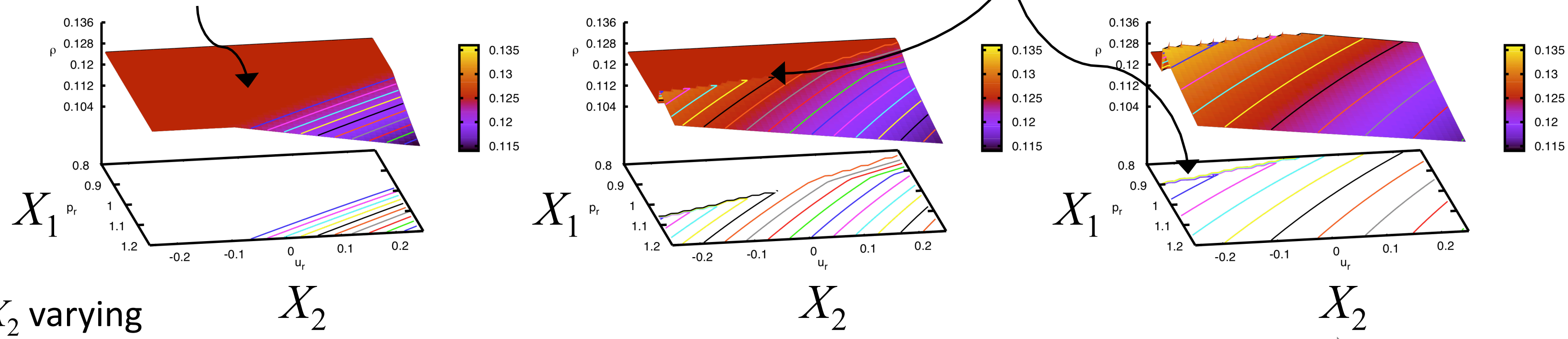


Output surface slice for the exact solution

The $\underline{\text{exact}}$ response surface for Y_2, the density at x = 1.16, is quite different.

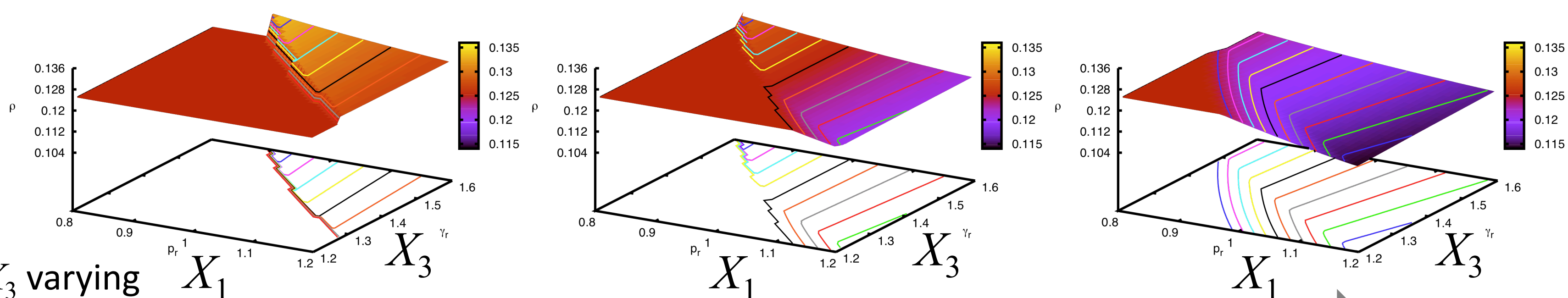
Flat regions mean that no waves have reached here

Sharp "cliffs" mean shocks



X_1, X_2 varying
X_3 fixed

X_3 increasing



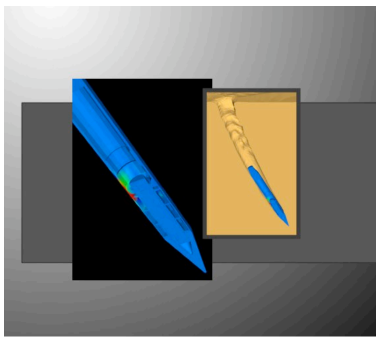
X_1, X_3 varying
X_2 fixed

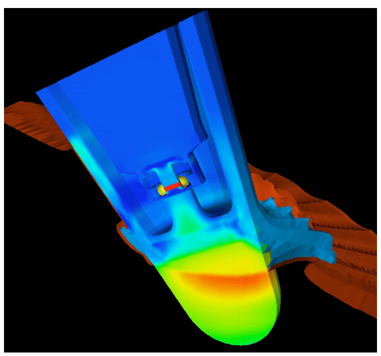
X_2 increasing

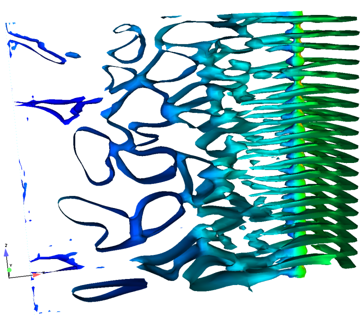
Sandia National Laboratories

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

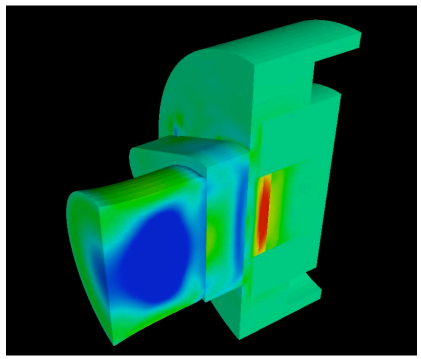
Shock and Multi-physics

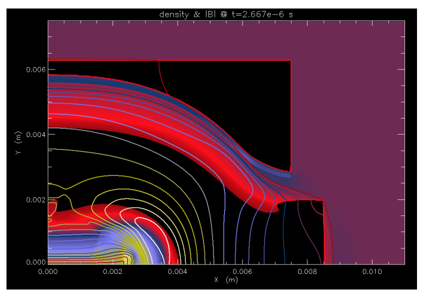


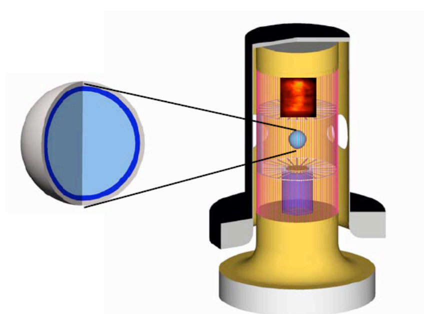




HEDP Theory and ICF Target Design





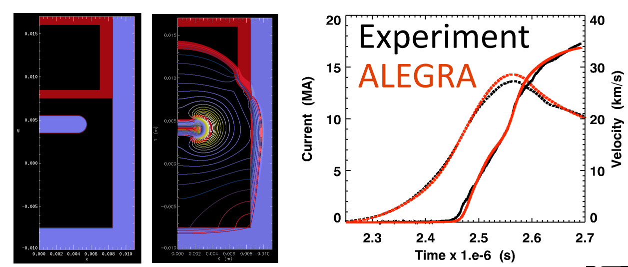


Overview

- 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

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



Isentropic Compression: Magnetic Flyer Prediction vs. Exp't

Sandia National Laboratories

The underlying equations in ALEGRA are related to hyperbolic conservation laws.

- The fundamental equations are statements of conservation laws:

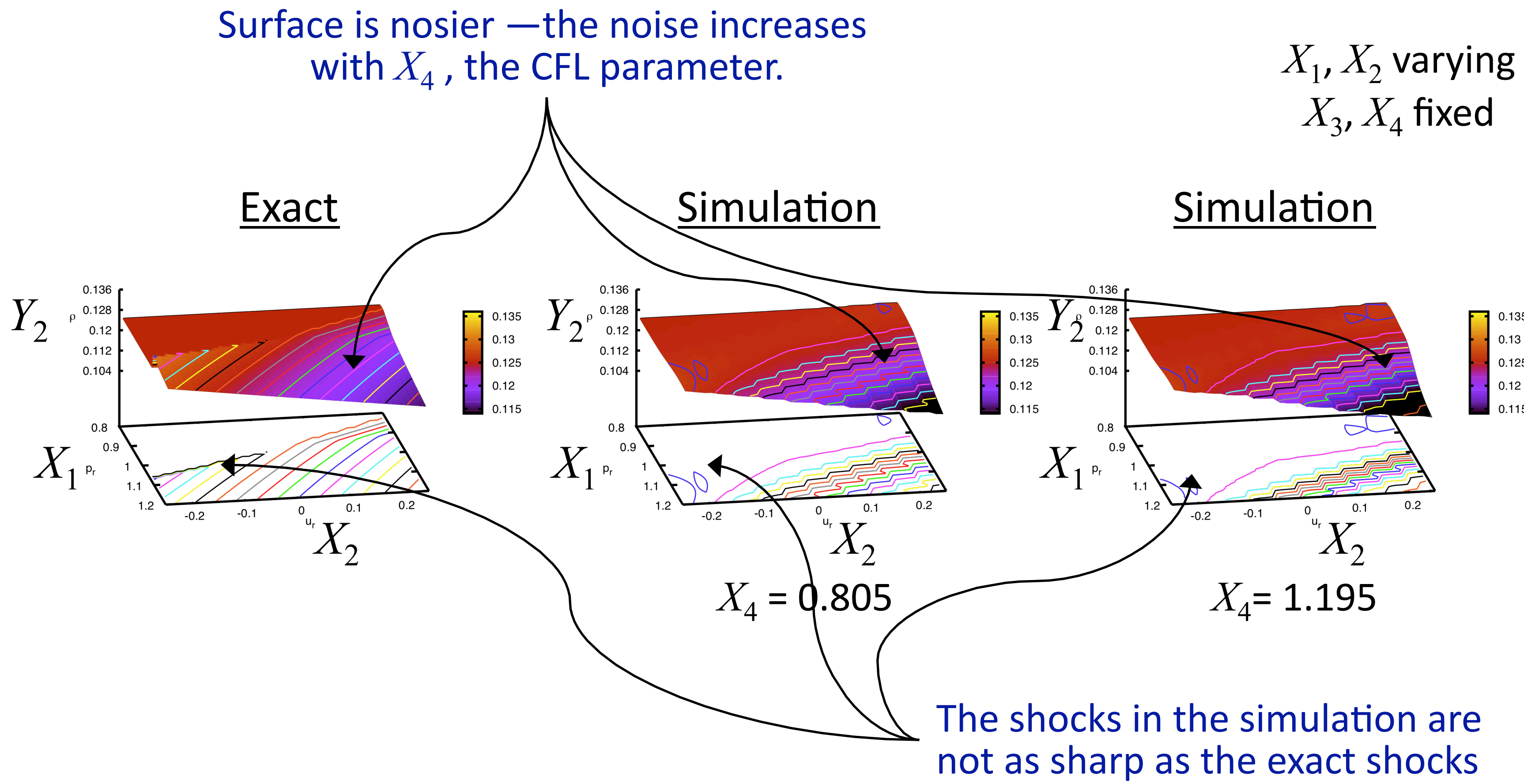
$$\frac{\partial U}{\partial t} + \mathrm{div}\, f(U) = S(U) \qquad x \in \Omega \subset \Re^3, \; t \geq 0$$

State — U; Flux function — $f(U)$; Source term — $S(U)$

  - Depending on the physics modeled, the state U may include, e.g.:
    - Internal state variables from material strength models
    - Magnetic field quantities for MHD simulations
  - These are discretized on a hexahedral mesh in the Arbitrary Lagrangian-Eulerian framework, amenable to general meshing and remapping.
- The gas dynamics equations are the "simplest" nonlinear physics equations that form a basic part of the full set of models in ALEGRA.
- This study is a prototype for the future analysis of problems with more complicated physics.

Sandia National Laboratories

The simulation response for Y_2, the density at x = 1.16, is different from the exact response.



- For this problem, most simulation response surfaces differ only slightly from the exact response surfaces.

We show results for estimators of the main (S) and total (T) sensitivity indices for several methods.

Meta-models
- DACE 256 Gaussian process approach, 256 samples
- ACOSSO 256 adaptive smoothing spline, 256 samples
- SDP 256 non-parametric smoothing, 256 samples

Analytic VBD
- PCE6 1296 analytic VBD, 6th-order, uniform distr., 1296 samples
- JRC 196k 196k sample, Sobol´/Saltelli estimates *The usual "gold standard"*
- PCE4 256 analytic VBD, 4th-order, uniform distr., 256 samples

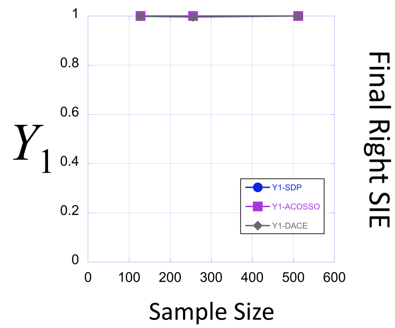
LHS Sampling
- LHS 60000 6.e+4 samples, Latin Hypercube Sampling VBD
- LHS 6000 6.e+3 samples, Latin Hypercube Sampling VBD

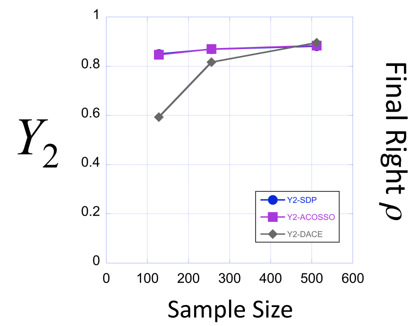
Full Factorial
- A-EXACT 160k 1.60e+5 (=20^4) ALEGRA, "full factorial" VBD
- A-EXACT 2.56M 2.56e+6 (=40^4) ALEGRA, "full factorial" VBD
- R-EXACT 160k 1.60e+5 Riemann (exact), "full factorial" VBD
- R-EXACT-2.56M 2.56e+6 Riemann (exact), "full factorial" VBD

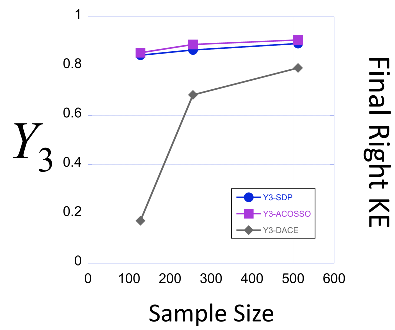
Sandia National Laboratories

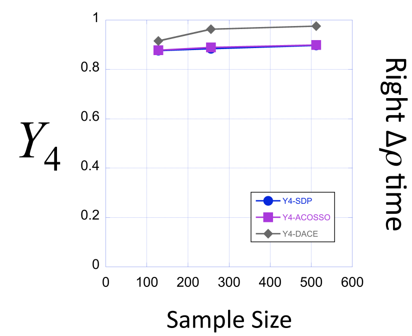
Comparison of R^2 for different meta-models with different sample sizes gives a measure of the goodness-of-fit.

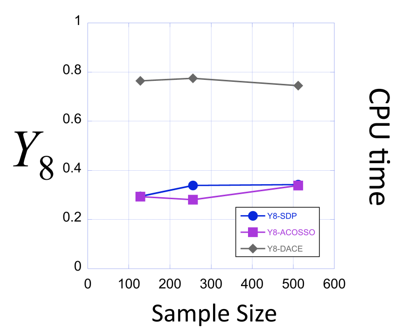
- The R^2 statistic is plotted for SDP, ACOSSO, and DACE (GP) emulators built with sample sizes: N=128, 256, 512.
- The goodness-of-fit clearly varies with the output:
  - Y_1 is very well fit
  - Y_2, Y_4 are reasonably well fit
  - Y_3 is reasonable with SDP, ACOSSO, but not so well with DACE (GP)
  - Y_8 is fit consistently better with DACE than the consistently poor fit with SDP and ACOSSO







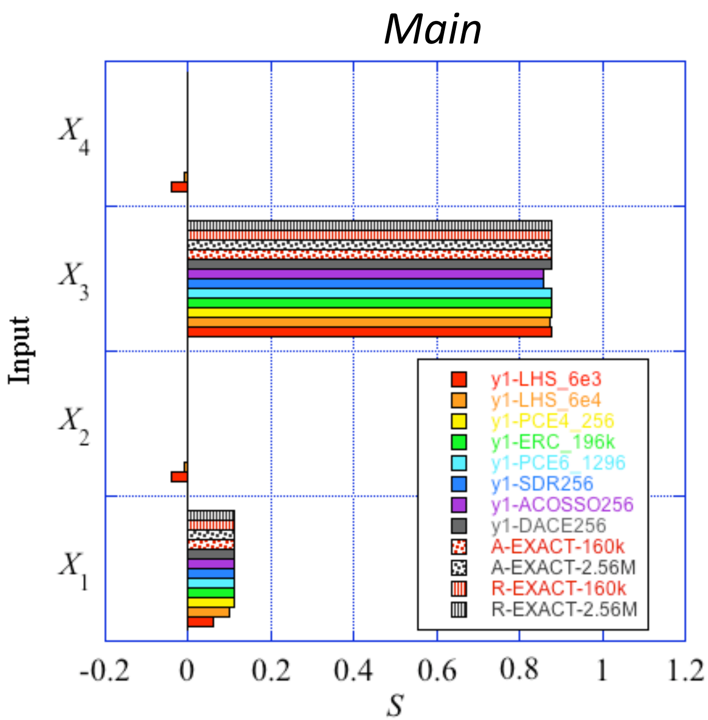


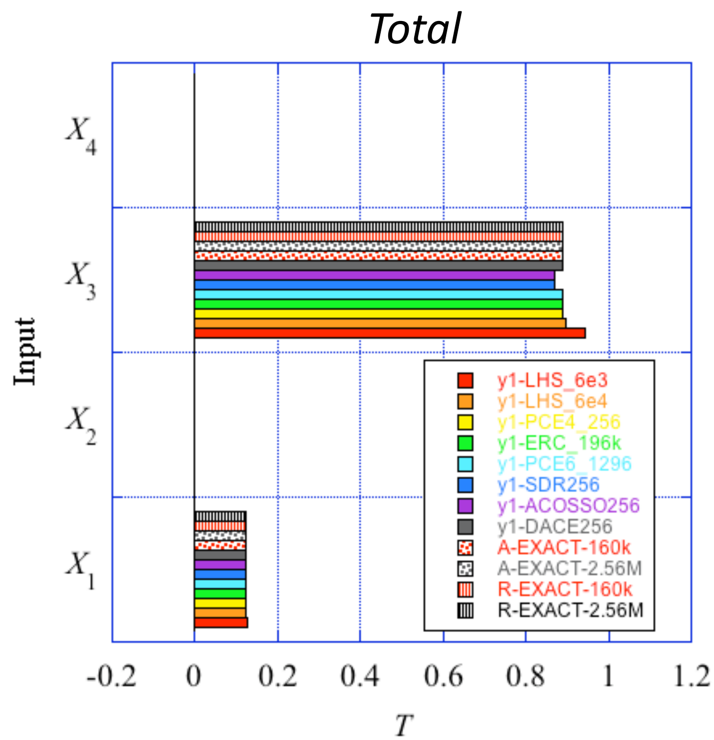


Sandia National Laboratories

The sensitivity indices S and T for Y_1 perform similarly for all approaches.

- As anticipated, Y_1 (SIE) depends strongly on X_1 (p_R) and X_3 (γ_R)
- Sampling, meta-model, and "exact" results are all consistent.

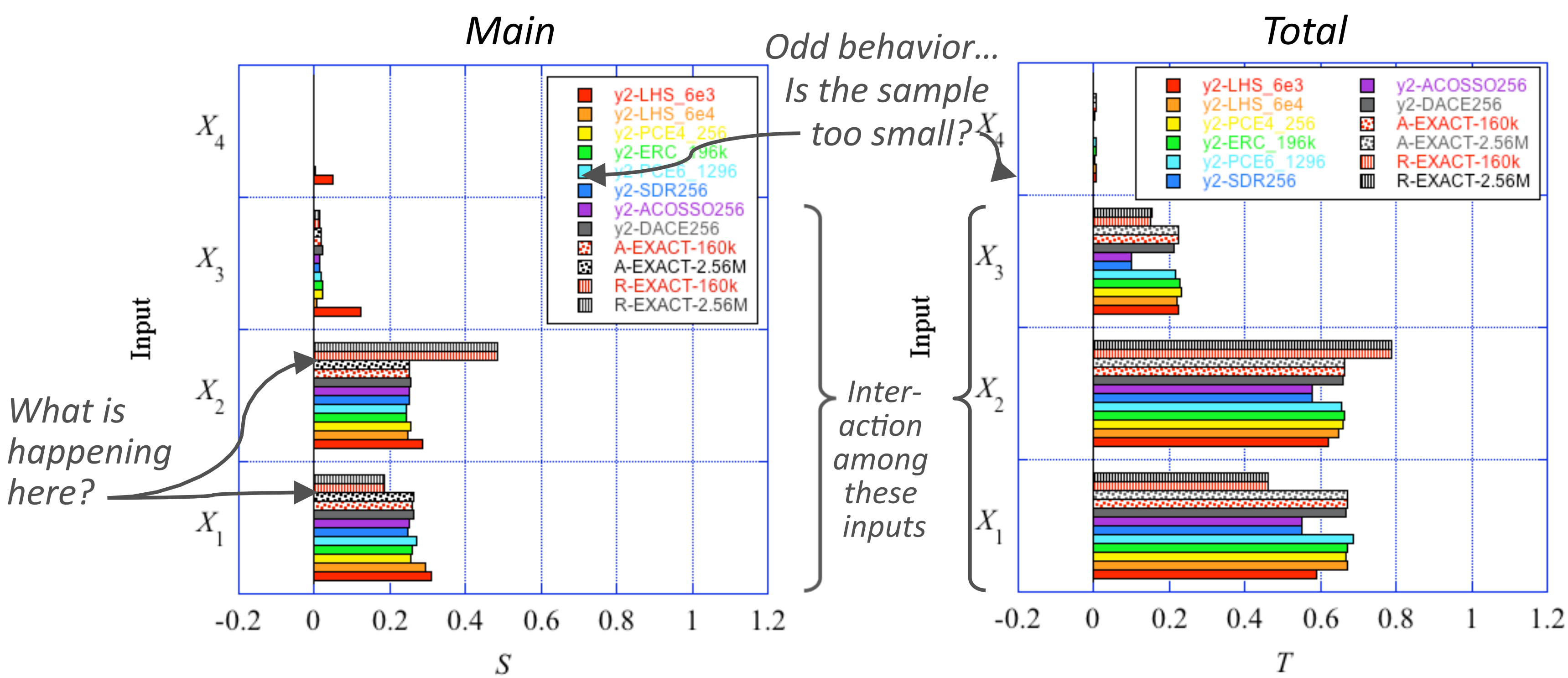




LHS 6000	PCE4 256	PCE6 1296	ACOSSO 256	A-EXACT 160k	R-EXACT 160k
LHS 60000	JRC 196k	SDP 256	DACE 256	A-EXACT-2.56M	R-EXACT-2.56M

The sensitivity indices for Y_2 have some unusual features.

- For Y_2 (final right ρ), LHS has different ranking, particularly for 6.e+3 samples and esp. wrt X_3 (γ_R) and X_4 (CFL).

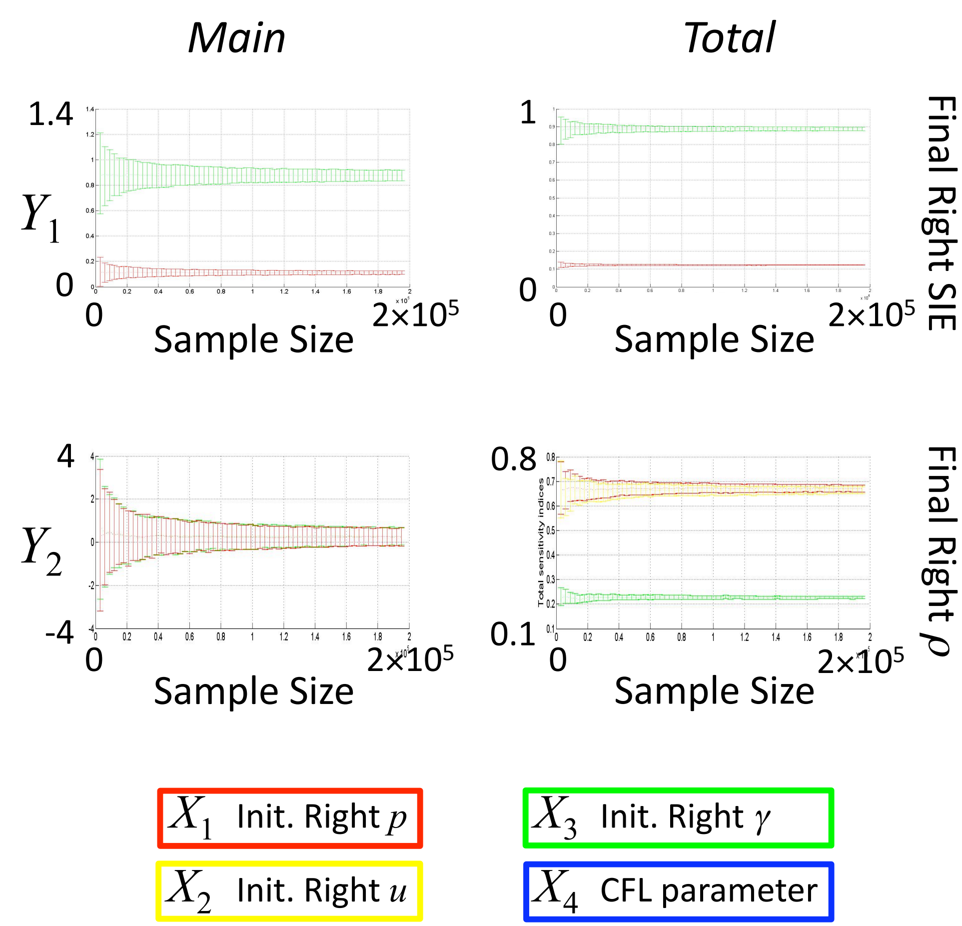


LHS 6000 | PCE4 256 | PCE6 1296 | ACOSSO 256 | A-EXACT 160k | R-EXACT 160k

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

Estimators of the main and total sensitivity indices§ converge under quasi-random sampling.

- Confidence intervals were calculated with a bootstrap technique.*
- Confidence intervals *decrease* with increasing number of model runs.
- The lower/upper bounds of the main indices are wider than those of the total indices.
- The estimator of the main indices to have a larger variance than the estimator of the total indices.

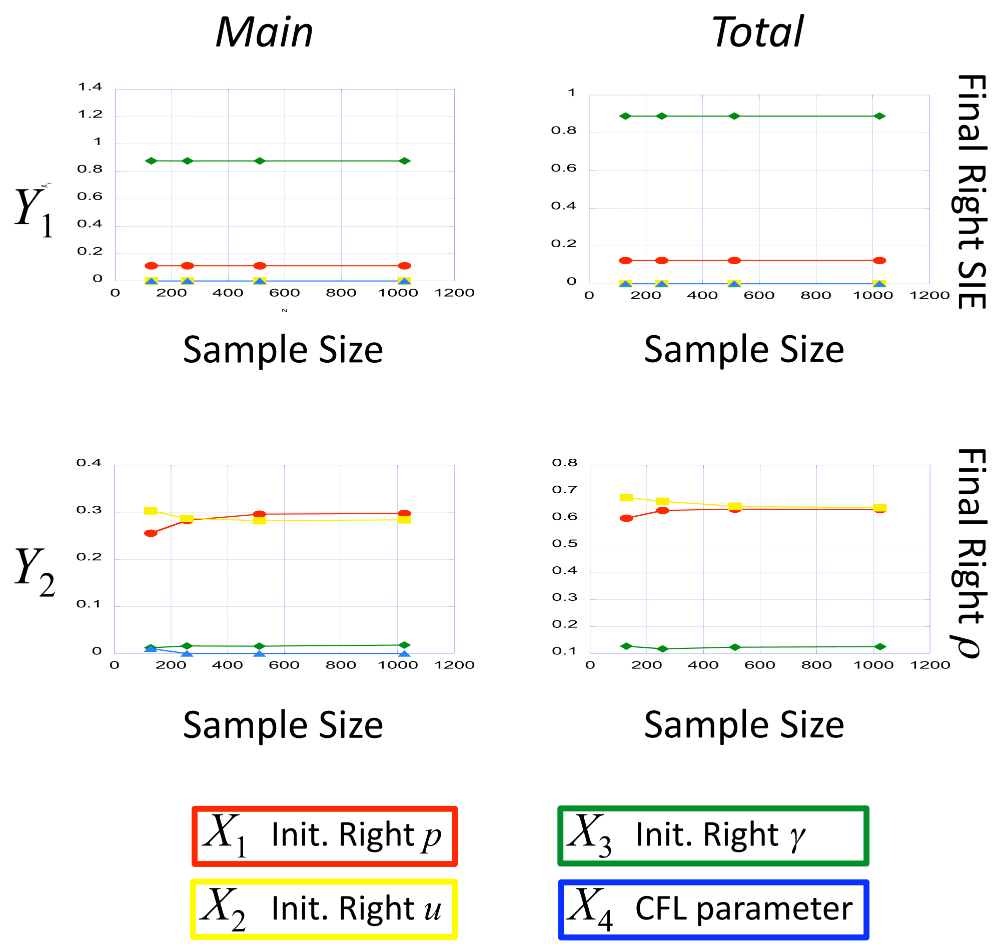


§ Saltelli, A., P. Annoni, I. Azzini, F. Campolongo, M. Ratto, S. Tarantola, "Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index," *Comp. Physics Comm.*, **181**, 259–270 (2010).

* G.E.B. Archer, A. Saltelli, I.M. Sobol´, "Sensitivity Measures, ANOVA-Like Techniques and the Use of Bootstrap," *J. Statist. Comput. Simul.*, **58**, pp. 99–120 (1997).

The main and total indices for the SDP meta-model converge with sample size.

- Meta-model results are for SDP + Sobol´ estimators built with sample sizes: N=128, 256, 512, 1024.
- Sobol´ indices are calculated with the meta-model at a set of "untried" points, i.e., points not used to build the meta-model.
- Both main and total indices are well-behaved with respect to convergence.
- The indices from N=256 are robust to further refinement.



We have some answers to our questions…

- *Do these approaches give consistent results, e.g., for rankings?*
  - In general, the different meta-models are consistent, both in ranking and magnitude, particularly for main effects (less so for total effects).
- *Do these results vary for the different outputs?*
  - "Well-behaved" outputs (e.g., Y_1 and Y_3) are quite consistent.
- *How to these results depend on the different inputs?*
  - "Well-behaved" inputs (e.g., X_1, X_2) follow the above pattern.
  - Other inputs (X_3, X_4) show more variation for SDP and ACOSSO.
  - Correct index values can be more challenging to properly calculate when there are significant interactions among the inputs (e.g., Y_2)

Sandia National Laboratories

We have some answers to our questions…

- *Do these results "converge"?*
  - Yes (empirically): more samples → the results "settle down"
  - No: the "converged" value might differ from the *exact* value.
- *How to sampling and meta-model results compare?*
  - In general, these two methods give comparable results.
- *Can we distinguish among different meta-models?*
  - The actual numbers varied slightly, but the rankings are robust.
- *How to exact solution results compare to ALEGRA results?*
  - "Well-behaved" inputs (e.g., X_1, X_2) follow the above pattern.

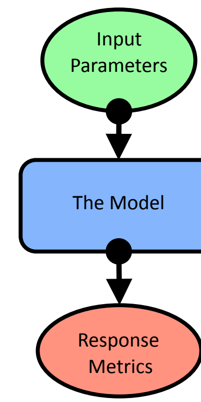
- LHS 6000
- LHS 60000
- PCE4 256
- JRC 196k
- PCE6 1296
- SDP 256
- ACOSSO 256
- DACE 256

Sandia National Laboratories

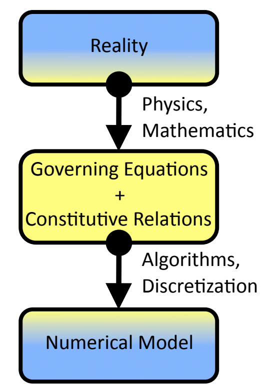
Summary: What did we talk about?

- Sensitivity Analysis :
  - Sobol'/Saltelli estimators of indices from quasi-random sampling DAKOTA JRC EUROPEAN COMMISSION ipsc
  - Sensitivity analysis using meta-models
    - PCE
    - SDP
    - ACOSSO
    - DACE (GP)

  PCE: DAKOTA — SDP, ACOSSO, DACE (GP): JRC EUROPEAN COMMISSION ipsc



- The Application:
  - The specific problem considered—and why
  - Inputs, outputs, and what we expected



- Computer Simulations:
  - The sensitivity analysis of the simulation model does not always match that of the exact model

Conclusions: What we determined

- We considered real-physics test problem, with an exact sol'n.
- The response surfaces for computed and exact solutions were compared and exhibited discontinuous behavior.
- Monte Carlo sampling gave bounded convergence for standard sensitivity measures.
- All meta-models gave consistent main effects index values.
- Greater variability was seen for some outputs with both "small" and "large" LHS-based indices.
- Differences between the computational model and the exact model were observed.
- This study led to improvements in DAKOTA algorithms.

Sandia National Laboratories