Tail Fragility as a Tool for Model Confidence

Ryan G. McClarren
University of Notre Dame

HOW DO WE ASSESS A COMPUTATIONAL METHOD?

QUAM BENE NON QUANTUM

- Most of us are raised on error analysis that speaks of convergence rates.
  - These convergence rates measure the theoretical error as some discretization parameter tends to 0.
- Important to keep in mind that these are rates and not absolute measurements
  - On a given problem, with a given mesh, etc., a first-order method may have less numerical error than a fourth-order one.
- We also might use efficiency or cost metrics as well
  - Memory usage, operations per degree of freedom, strong or weak scalability, etc.

HOW DO USERS ASSESS NUMERICAL METHODS?

QUAM BENE VIVAS REFERRE, NON QUAM DIU

- Time to solution for a given problem
- Lack of obvious errors that clearly violate physical principles
- Lack of too much sensitivity to small changes in problem settings.

A CODE USER'S WORKFLOW DOES NOT RESEMBLE A CONVERGENCE STUDY

A RECIPE FOR SUCCESSFUL COMPUTATION (?)

- Try to solve the problem with a refined geometry and numerical parameters (good mesh, tight tolerances on time step control, and refined in every sense) as advised by the code developers and computational scientists.
- Find out that it takes too long to get the solution, so coarsen something (or everything) and try again.
- Repeat until one gets a solution in a reasonable amount of time.
- Use the settings/mesh to run a design or parameter study.
- Justify any possible errors (due to numerics, model equations, etc.) on a post hoc basis.
- Complain about code developers (this can be done at every step).

WHAT THE DEVELOPERS AND USERS CARE ABOUT ARE NOT EXACTLY THE SAME

MEASURING LOCAL ERROR CONVERGENCE RATES IS NOT ENOUGH

QUI TOTUM VULT TOTUM PERDIT

- The fact that higher-order methods typically are more efficient in a measure of computational cost per unit error is often postulated as a reason to use these methods.

- Quantifying the local error is what most of us are trained to do (and how many methods are derived).

- Most “real” problems have solutions that violate the assumptions needed for high-order error convergence (non-smooth solutions, shocks, etc.).

- Higher-order methods also imply that the error increases at a faster rate when the mesh or other parameters are coarsened.

LOCAL ERROR CONVERGENCE IS NOT WHAT USERS WANT

- Someone using simulation to design a component has specific design criteria or Quantities of Interest (QoI).
  - It is those QoIs that the error is important in.
- Also, there are uncertainties in the problem inputs that will effect the QoIs.
- When a user cares about the uncertainty in the QoI, the accuracy of the uncertainty estimates will also be important.
- It does not immediately follow that a numerical method that has good behavior with respect to local error will give adequate estimates of uncertainty.

A THOUGHTFUL USERS DESIDERATA

- A user of code will, one hopes, be aware of the fact that a code solution is not the truth.

- Under these conditions a user should desire a method that yields estimates of QoIs and their uncertainty that are not sensitive to the particular choices the user has to make.

    - A small change in the discretization parameters should not lead to a large change in the QoIs and their associated uncertainties.

    - A small change in the assumptions of the analyst should not have a large change in the output.

A REGULATOR OR DECISION MAKER WORRIES ABOUT THE EXTREME CASES

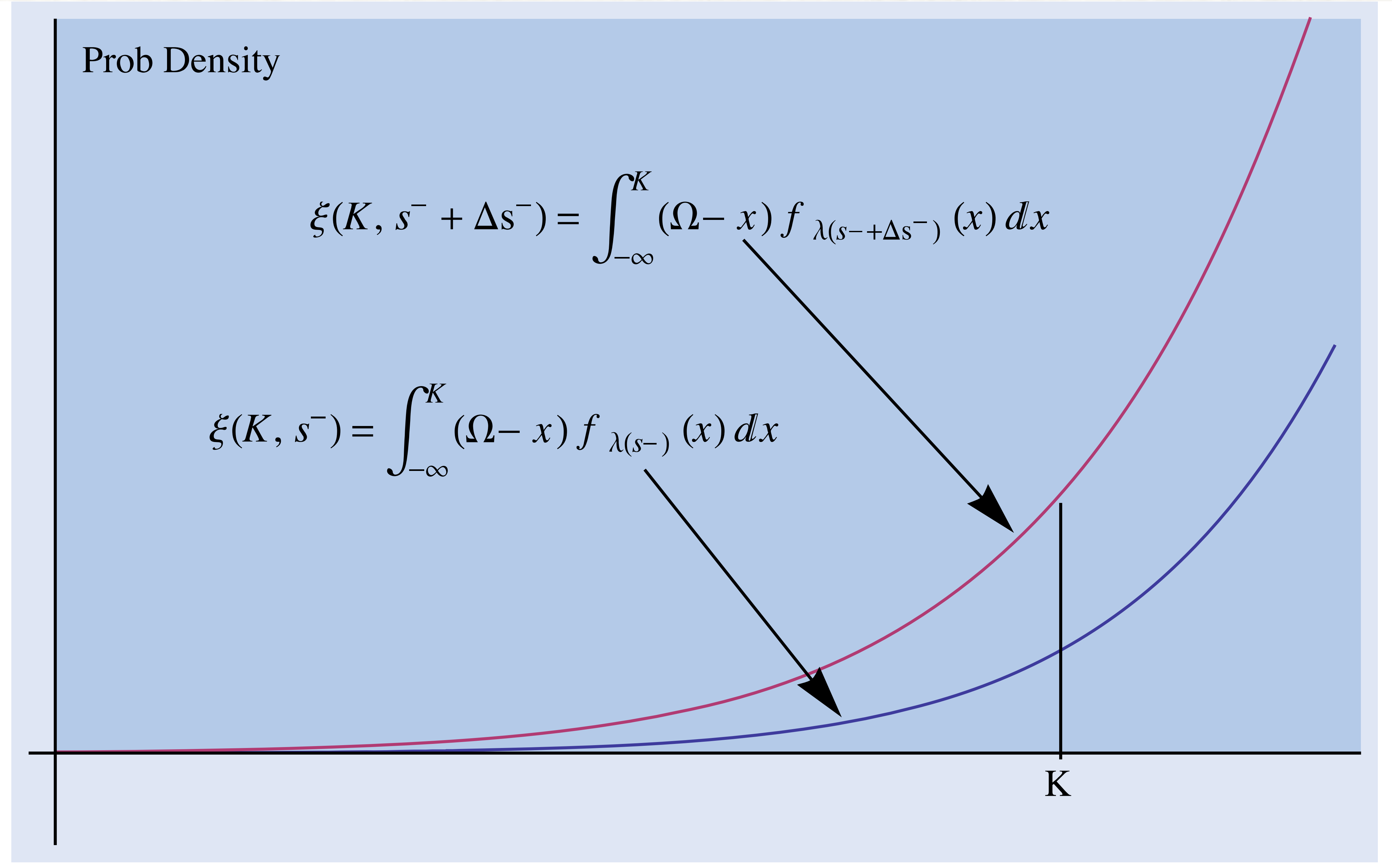
- When using simulation to inform a decision about safety, it is not the nominal case that is of interest, rather what happens when things go wrong.
- What is the probability of failure?
- The sensitivity of this parameter to inputs and discretization choices is important to making the solutions credible.
- The choice of physical model has a crucial role in this.
  - Does the model represent reality at these extremes?
- In 2008, most mortgage risk models assumed that the risk of default on a mortgage was independent of the risk in other loans.
  - This is true in normal times, but when everyone in a neighborhood defaults, the risks are highly correlated.

HOW FRAGILE ARE UNCERTAINTY ANALYSES TO ASSUMPTIONS

- When one does perform an uncertainty analysis there are necessarily assumptions made in the process.
  - Type of input distributions, distribution parameters (e.g., variance), tail behavior etc.
- These assumptions lead to epistemic uncertainty and their implications are not distributional.
- Taleb and Douady (2013) use these uncertainties in input distributions to define the fragility of a uncertainty analysis:
  - For instance, how much does the probability of failure for the system increase with an increase in the variance in an input parameter?
  - Is the second derivative of the increase positive or negative?
- Different mathematical models will have different behaviors in this sense.

THE FRAGILITY MEASURE OF TALEB AND DOUADY

- How much does the integral of the lower tail of the probability density change when a parameter of the probability distribution changes.
- K is the point below which something bad happens.
- $\lambda(s^-)$ is a parameter in the probability distribution set to make certain amount of mass s^- in the left tail.
- Ω is the center of the distribution.



NEW METRICS TO BENCHMARK NUMERICAL METHODS AND PHYSICS MODELS ARE NEEDED

- If we care about uncertainty and nominal values of a quantity of interest, we need to think about how the choices made when running the code affect the analysis.
- The solution verification community has worked on how to quantify numerical error, but the question of model error is much harder.
  - How does my choice of subgrid model A over model B affect my results, and which one is more correct?
- Propagating the numerical and model error to the uncertainty in the final answer is not straightforward.
- I'll present some ideas for metrics that seek to measure how choices in the solution process affect the results.

WE CONSIDER A QOI THAT IS A FUNCTION OF DISCRETIZATION AND RANDOM VARIABLES

- Consider a QoI, $Q(\lambda, \xi)$, where λ are discretization parameters and ξ are random variables.

- Two measures of uncertainty that we are interested in are

  – The variance in the QoI due to the random inputs,

$$\mathrm{V}(Q)(\lambda) = \int_{-\infty}^{\infty} Q(\lambda, \xi)^2 p(\xi)\, d\xi - \left(\int_{-\infty}^{\infty} Q(\lambda, \xi) p(\xi)\, d\xi \right)^2$$

  – The probability of failure, P_{fail}, i.e., the probability that Q exceeds some threshold,

$$P_{\text{fail}} = 1 - F_Q(Q_U, \lambda),$$

  where F_Q is the CDF of $Q(\lambda, \xi)$ and Q_U is the failure point.

- Both of these quantities functions of the mesh parameters, λ.

THE VARIANCE IS EASIER TO ANALYZE THAN THE PROBABILITY OF FAILURE

- The sensitivity of these measures to the discretization parameters can be found by differentiating with respect to a single parameter λ_i.
- For the variance this is

$$\begin{aligned}\frac{\partial V}{\partial \lambda_i} &= \int_{-\infty}^{\infty} 2Q(\lambda,\xi)\frac{\partial Q}{\partial \lambda_i} p(\xi)\, d\xi - 2\int_{-\infty}^{\infty} Q(\lambda,\xi) p(\xi)\, d\xi \int_{-\infty}^{\infty} \frac{\partial Q}{\partial \lambda_i} p(\xi)\, d\xi \\ &= 2\mathrm{E}\left[Q\frac{\partial Q}{\partial \lambda_i}\right] - 2\mathrm{E}[Q]\mathrm{E}\left[\frac{\partial Q}{\partial \lambda_i}\right]\end{aligned}$$

  where $\mathrm{E}[g]$ is the expected value of a function g with respect to ξ.
- The sensitivity of the variance to a mesh parameter is expressed in terms of expected values involving the QoI and its sensitivity.
- The probability of failure's sensitivity is

$$\frac{\partial P_{\text{fail}}}{\partial \lambda_i} = -\frac{\partial F}{\partial \lambda_i}(Q_U, \lambda).$$

- This quantity is what I called tail fragility in the abstract.
- Unfortunately, there is a not a simple way to relate this to expected values. The best we might be able to do is a sensitivity of the result from a Monte Carlo study.

A LARGE VALUE FOR EITHER OF THESE SENSITIVITIES IS A RED FLAG

- If the variance or the probability of failure is sensitive to the discretization parameters, it is hard to have faith in the results.
- The nominal values may be less useful than the relative values.
- Even if these sensitivities are not rigorously calculated, they should be estimated.
- If I were a making a critical decision, especially one with large consequences, I would ask an analyst for these numbers.
  - If the analyst has no clue as to these numbers, how can I have confidence.
- The discretization parameters in the above could also be choices in the underlying uncertainty (e.g., range or variance of the uncertain inputs).

THE SENSITIVITY OF THE VARIANCE CAN BE EXPRESSED IN TERMS OF A POLYNOMIAL EXPANSION

- If we expand the QoI in a polynomial chaos expansion,

$$Q(\xi, \lambda) = \sum_{\ell=0}^{\infty} c_\ell(\lambda)\phi_\ell(\xi),$$

where $\phi_\ell(\xi)$ is an orthonormal basis function and

$$c_\ell(\lambda) = \int_{-\infty}^{\infty} Q(\xi, \lambda)\phi_\ell(\xi)p(\xi)\, d\xi.$$

- The sensitivity to the variance is then

$$\frac{\partial V}{\partial \lambda_i} = 2\sum_{\ell=1}^{\infty} c_\ell(\lambda)\frac{\partial c_\ell}{\partial \lambda_i}.$$

- This indicates that when we compute the expansion coefficients, we also need to compute

$$\frac{\partial c_\ell}{\partial \lambda_i} = \int_{-\infty}^{\infty} \frac{\partial Q}{\partial \lambda_i}(\xi, \lambda)\phi_\ell(\xi)p(\xi)\, d\xi.$$

THE SENSITIVITY OF THE VARIANCE CAN BE EXPRESSED IN TERMS OF A POLYNOMIAL EXPANSION

- The sensitivity to the variance is then

$$\frac{\partial V}{\partial \lambda_i} = 2 \sum_{\ell=1}^{\infty} c_\ell(\lambda) \frac{\partial c_\ell}{\partial \lambda_i}.$$

- This indicates that when we compute the expansion coefficients, we also need to compute

$$\frac{\partial c_\ell}{\partial \lambda_i} = \int_{-\infty}^{\infty} \frac{\partial Q}{\partial \lambda_i}(\xi, \lambda) \phi_\ell(\xi) p(\xi) \, d\xi.$$

- If the sensitivity of the QoI is expanded

$$\frac{\partial Q}{\partial \lambda_i} = \sum_{\ell=0}^{\infty} \frac{\partial c_\ell}{\partial \lambda_i} \phi_\ell(\xi),$$

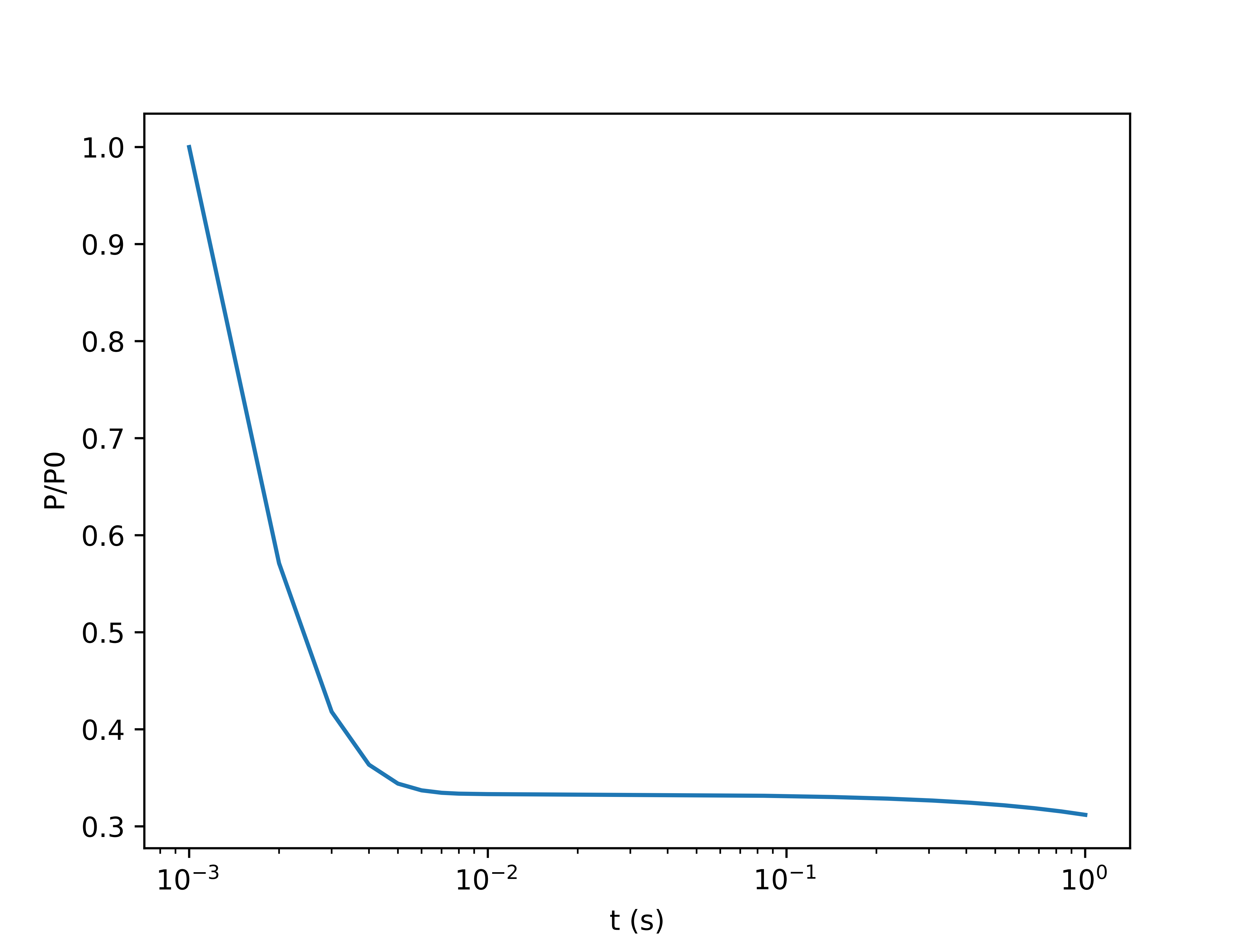
  then we have the information we need to compute $\partial V / \partial \lambda_i$.

- Note that the sensitivity of the variance is not the same as the variance of the sensitivity.

- This result gives further evidence that adjoint calculations can be useful (the sensitivity information to any parameter is available).

EXAMPLE CALCULATION: POINT REACTOR KINETICS

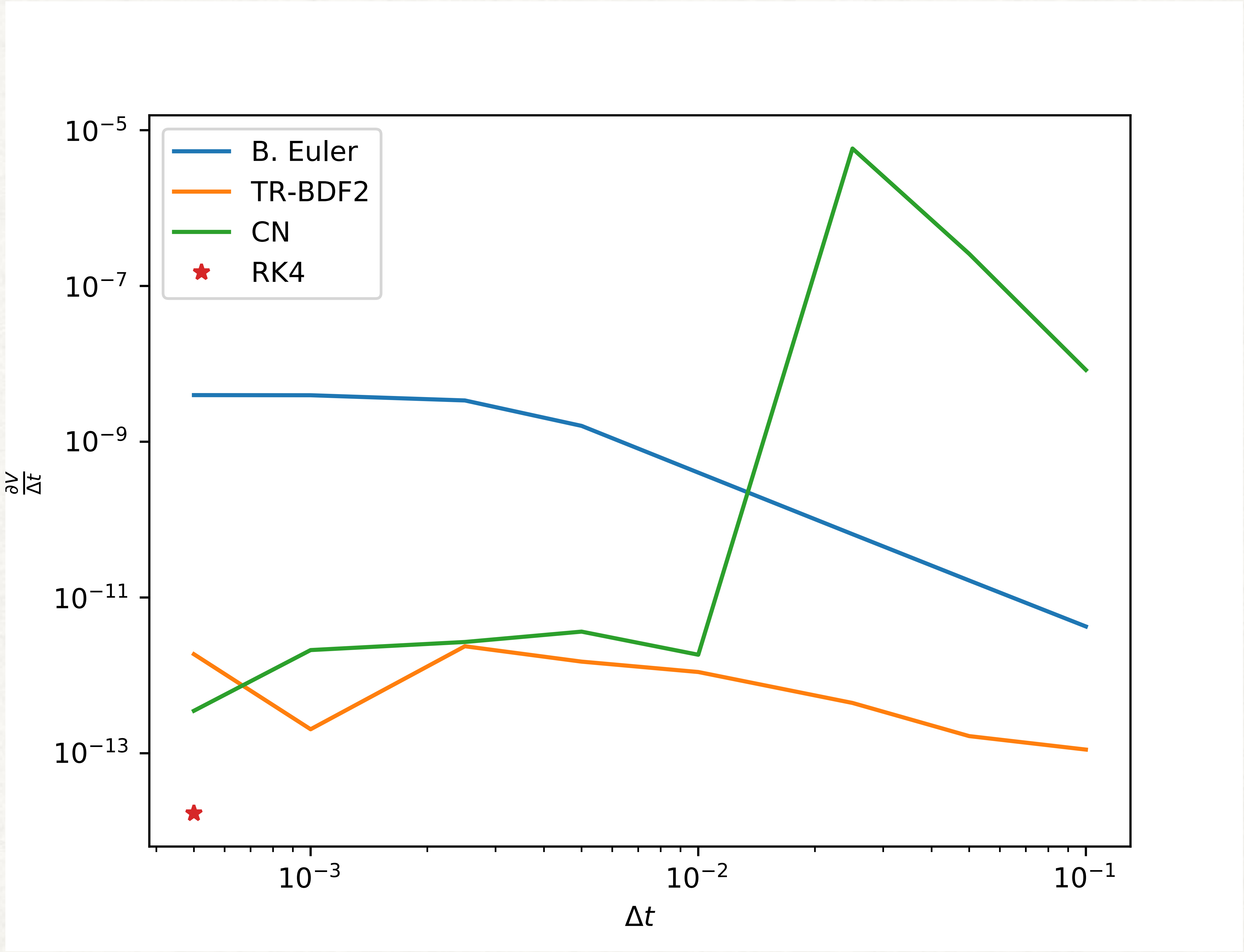
- Point Reactor Kinetics model the power of a nuclear reactor using a system of ODEs.
- We will look at a simple system of a reactor shutdown by inserting a control rod at time 0 (scram).
- We prescribe that system fail occurs if the power has not dropped by a factor 0.325 in 1 second.
- The control rod worth is a normal random variable with mean -2 and standard deviation 0.05.
- We will look at how our sensitivity measures change with a change in the time step size used, and with changes in the standard deviation.



Rod Drop solution with RK4

WITH A LARGER TIME STEP CERTAIN METHODS HAVE A LARGER SENSITIVITY TO TIME STEP SIZE

- We tested backward Euler, Crank-Nicolson, TR-BDF2, and Runge-Kutta 4, to integrate the ODEs with a fixed time step.
  - We perturb the time step size to compute the derivative of the variance and probability of failure to the step size.
- For large time steps, Crank-Nicolson displayed a larger sensitivity to the estimate variance in the power at time 1 with respect to the time step size.
- When the time steps get small, backward Euler was more sensitive in the variance.
- The probability of failure was not sensitive to the time step size.



THE PROBABILITY OF FAILURE IS CONVEX TO THE ESTIMATED STANDARD DEVIATION IN THE INPUTS

- When we increase or decrease the standard deviation of the control rod worth by 10% we find:
  - As the SD goes up, the probability of failure goes up.
  - Additionally, the second derivative is positive, implying a larger change will have an even larger effect on the probability of failure.
- The variance in the estimate has a positive first-order sensitivity, but a negative second derivative.

$$\frac{\partial P_{\text{fail}}}{\partial \sigma} \approx 1.48$$

$$\frac{\partial^2 P_{\text{fail}}}{\partial \sigma^2} \approx 104$$

$$\frac{\partial V}{\partial \sigma} \approx 0.001219$$

$$\frac{\partial^2 V}{\partial \sigma^2} \approx -1.852$$

THE MODELS, NUMERICS, AND ASSUMPTIONS ALL MATTER

- When using simulations to predict system behavior, we must be conscious of how our assumptions affect the outcome.
- I tried to argue that we should be measuring and reporting this.
- I did not talk about mathematical model selection. There could be cases where this has a large impact on the sensitivity of the analysis to assumptions.
  - Simpler models may be less fragile in the sense that a “bad” solution to them has a smaller effect on the results of an analysis than a complex model.
    - Diffusion versus Transport in radiative transfer
    - Helmholtz versus Maxwell
- Asking the hard questions as consumers of results is an important first step.