Department of Aerospace & Mechanical Engineering, University of Notre Dame

That's an interesting idea

Data driven models, compressed sensing, and other outré tools for nuclear applications.

Ryan G. McClarren
rmcclarr@nd.edu

Young Members Research Achievement Award Plenary
M&C 2019

August 26, 2019

First, a few words of thanks.

- Many people have influenced my research career
- Graduate School and Beyond
  - James Holloway (my PhD. advisor)
  - Tom Brunner (research, coding, and physics guide)
- Postdoc at LANL
  - Todd Urbatsch (implicit Monte Carlo)
  - Rob Lowrie (radiation hydrodynamics)
  - Jeff Densmore (implicit Monte Carlo and optimism)
- My Texas A&M Colleagues
  - Marvin Adams
  - Jim Morel
  - Jean Ragusa
- I owe any research success I may have had to my interactions with these people, and many others that I could list.

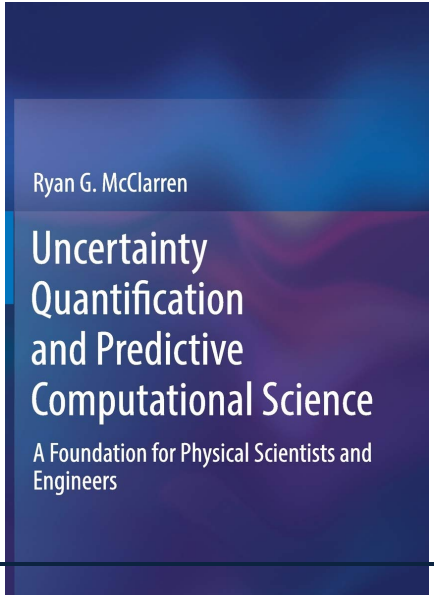
I have always tried to approach research by looking at problems in a unique way.

- In this talk I want to look at some potential avenues for innovation in mathematics and computation for nuclear applications.
  1. Nonlinear problems allow us to get more information from less data.
  2. The data produced from our simulations can be used to improve our computation.
- I cannot claim these ideas are 100% novel. I follow the applied mathematics, computational physics, and statistics literature *outside* the application areas I am interested in to draw inspiration.

Math & Comp is more than discretizations and numerical results.

- In this talk I want to look beyond the novel ways of solving the transport equation and think about the state of computation.
- Computers are getting faster and transport calculations are becoming standard in a variety of applications.
  - Low-order models subject to calibration are not essential to analyses.
- As a result numerical and transport model errors are becoming less important than errors arising from incorrectly specified data.
  - Cross-sections, problem specifications, etc. may cause a larger uncertainty than numerical error.
- Measuring these errors has been the focus in uncertainty quantification (UQ) for some time.
- An issue that looms that large is the number of uncertain parameters is generally much larger than we can deal with.

- Many of the ideas I will talk about, as applied to UQ, can be found in my book.
- Specifically, how to use compressed sensing ideas such as regularized regression to reduce the dimensionality of a UQ problem is discussed.
- Additionally, applying compressed sensing to polynomial chaos expansions is developed there as well.



We can use data to approximate operators and improve calculations.

- In scientific computing we are used to taking a known operator and making approximations to it.
- It is possible to use the action of an operator and use just the action of the operator to generate approximations to it.
  - This is the basis for many Krylov methods.
- We can use the action of radiation transport operators to
  - Estimate time eigenvalues present in a subcritical system, and
  - Compute the slowly converging modes in source iteration to accelerate convergence without the need for diffusion-based preconditioning [1]
  - Reduce the amount of memory needed to perform a simulation [2]

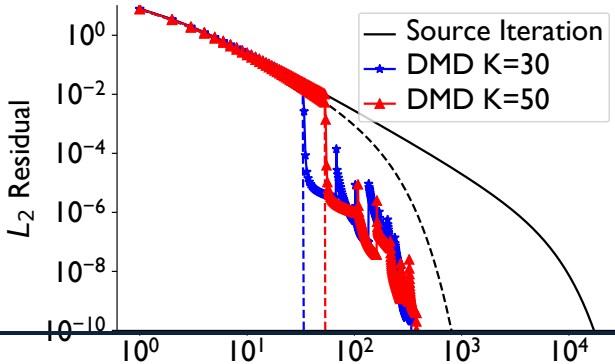
[1] See DMD talk Wednesday morning.

[2] See Low-Rank Talk Tuesday Morning in Transport Theory II.

- We consider a problem with vacuum boundaries, 1000 cells, unit domain length, with $c = 0.9999$ and

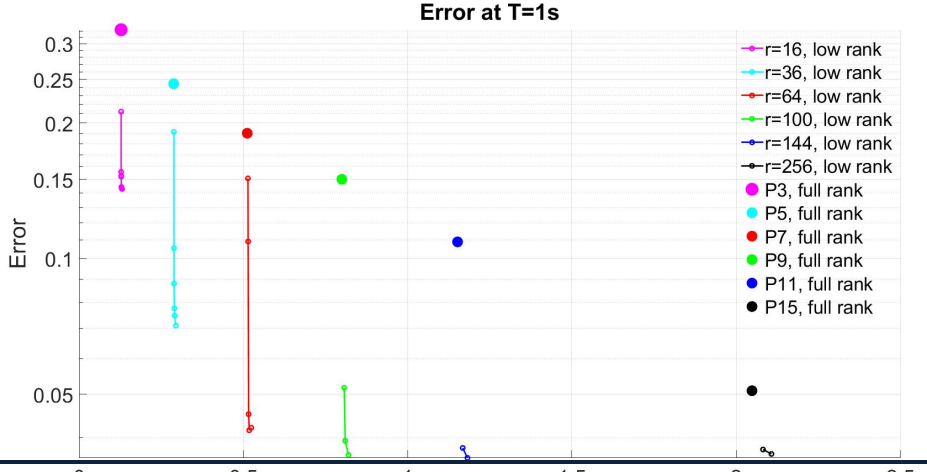
$$\sigma_t = \begin{cases} 2^p & \text{cell number odd} \\ 2^{-p} & \text{cell number even} \end{cases}.$$

- Below we see convergence for $p = 5$ (dashed) and $p = 8$ (solid), a factor of about 1000 and 6.5×10^4 between thick and thin cells, respectively.



Low-Rank Discretizations of the Transport Equation use much less memory. (Tuesday Morning w/Zhuogang Peng)

- On the linesource problem in 2-D Cartesian geometry, low-rank methods can use higher angular fidelity and obtain lower errors using less memory.



We begin with a sequence of vectors related by an operator.

- Consider a sequence of vectors $\{y_0, y_1, \ldots, y_K\}$ where $y_k \in \mathbb{R}^N$.
- The vectors are related by a potentially unknown linear operator of size $N \times N$, A, as

$$y_{k+1} = Ay_k.$$

- If we construct the $N \times K$ *data matrices* Y_+ and Y_-,

$$Y_+ = \begin{pmatrix} | & | & & | \\ y_1 & y_2 & \cdots & y_K \\ | & | & & | \end{pmatrix} \qquad Y_- = \begin{pmatrix} | & | & & | \\ y_0 & y_1 & \cdots & y_{K-1} \\ | & | & & | \end{pmatrix}$$

  we can write

$$Y_+ = AY_-.$$

- At this point we only need to know the data vectors y_k, they could come from a calculation, measurement, etc.
- As $K \to \infty$ we could hope to infer properties about A.

- We can rearrange the relationship between Y_+ and Y_- using the SVD:

$$Y_+ = AU\Sigma V^{\mathrm{T}} \qquad \rightarrow \qquad U^{\mathrm{T}}AU = U^{\mathrm{T}}Y_+V\Sigma^{-1}.$$

- Define $\tilde{A} = U^{\mathrm{T}}AU = U^{\mathrm{T}}Y_+V\Sigma^{-1}$. This is a rank K approximation to A.
- Using the approximate operator $\tilde{A}$, we can now find out information about A.
- The eigenvalues/vectors of $\tilde{A}$: $\tilde{A}w = \lambda w$, are used to define the dynamic modes of A:

$$\varphi = \frac{1}{\lambda}U^{\mathrm{T}}Y_+V\Sigma^{-1}w.$$

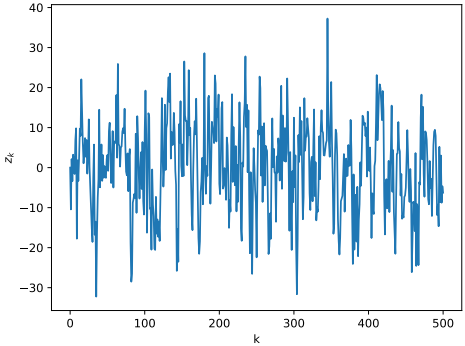
- The dynamic mode decomposition (DMD) of the data matrix Y_+ is then the decomposition of into vectors φ. The mode with the largest norm of λ is said to be the dominant mode.

Example of operator approximation for $N = 1$ demonstrates the efficacy.

- Consider the sequence

$$z_{k+1} = az_k + n_k,$$

  where $a = 0.5$, and $n_k \sim \mathcal{N}(0, 10^2)$.
- Using $K = 500$, we estimate $a = 0.506552$ from the data below.



DMD can be used to estimate eigenvalues of the matrix exponential.

- If we consider a sequence of vectors that are solutions to the system of differential equations,

$$\frac{\partial y}{\partial t} = Ay(t),$$

and are separated by a time, Δt, the relationship between vectors is

$$y^{n+1} = e^{A\Delta t} y^n.$$

- As before we can define Y_- and Y_+, compute the SVD of $Y_- = USV^*$, and approximate the matrix exponential:

$$U^{\mathrm{T}} e^{A\Delta t} U = U^{\mathrm{T}} Y_+ V \Sigma^{-1}.$$

- One can show the following:
  - The eigenvalues of $U^{\mathrm{T}} e^{A\Delta t} U$ are also eigenvalues of $e^{A\Delta t}$.
  - If α is an eigenvalue of A, then $e^{\alpha \Delta t}$ is an eigenvalue of $e^{A\Delta t}$.
  - The eigenvectors of A are eigenvectors of $e^{A\Delta t}$.

Current methods of estimating time-eigenvalues have issues for subcritical systems.

- In neutron transport for time-dependent problems α eigenvalues (also called time eigenvalues) are important quantities to understand system dynamics and safety.
  - These eigenvalues characterize the system evolution in terms of functions of the form $Ce^{\alpha t}$.
- The more well-known k-eigenvalue gives information about the long-term behavior of the system, but is less useful for diagnosing many experiments.
- Additionally, most alpha eigenvalue solvers have issues with subcritical problems due to “negative absorption”.
- For subcritical systems it has been shown that the rightmost eigenvalue in the complex plane is not necessarily meaningful.
  - There can be an eigenvalue with negative real part that is arbitrarily close to 0.
  - These eigenvalues correspond to the time scale of slow moving neutrons crossing the system.

The time evolution of the angular flux can be used to estimate alpha eigenvalues.

- If we compute $K+1$ time steps of size Δt using a transport solver, we will have the relation

$$\begin{pmatrix} | & | & & | \\ \psi_K & \psi_{K-1} & \cdots & \psi_1 \\ | & | & & | \end{pmatrix} = e^{A\Delta t} \begin{pmatrix} | & | & & | \\ \psi_{K-1} & \psi_{K-2} & \cdots & \psi_0 \\ | & | & & | \end{pmatrix}$$

  or $Y_+ = e^{A\Delta t} Y_-$ and we can take the SVD of Y_- as before.

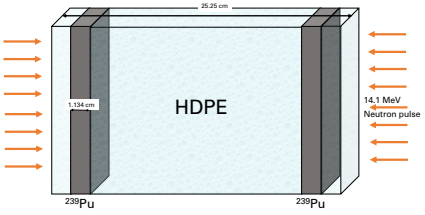
- Therefore, if we estimate the eigenvalues λ of the $K \times K$ matrix $U^{\mathrm{T}} e^{A\Delta t} U$, we can compute the alpha eigenvalues of the system as

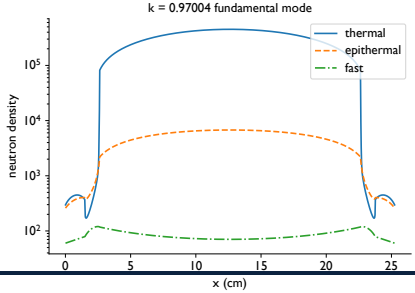
$$\alpha = \frac{\log \lambda}{\Delta t}.$$

- We do not need to a special eigenvalue solver to do this.

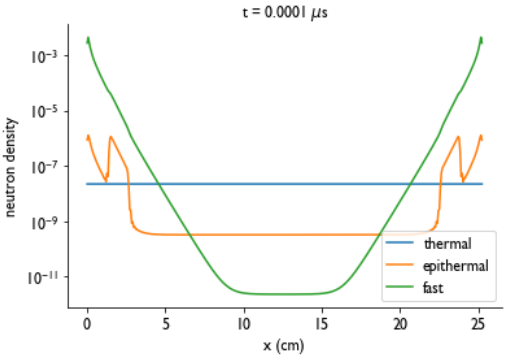
A heterogeneous, subcritical system will test this method on a nontrivial problem.

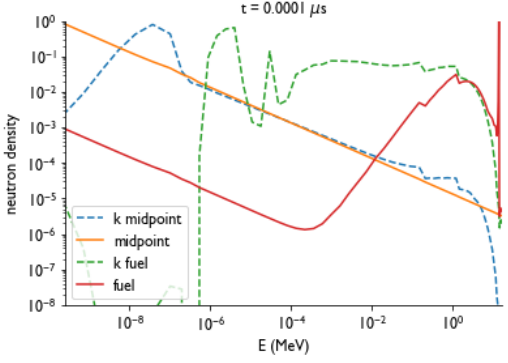
- We consider a slab-geometry problem where a plastic moderator is sandwiched between two slabs of plutonium with a small outer reflector.
- This system will have α eigenvalues associated with the time scale of slow neutrons crossing the moderator.
- The fundamental mode from a k-eigenvalue problem has many thermal neutrons in the middle of the problem.
- We consider the situation where DT neutrons enter the slab from both sides at time 0.
- Thermal: $E < 5$ eV, Fast: $E > 0.5MeV$





In 1 μs the spectrum does not approximate the fundamental k-eigenvalue mode.

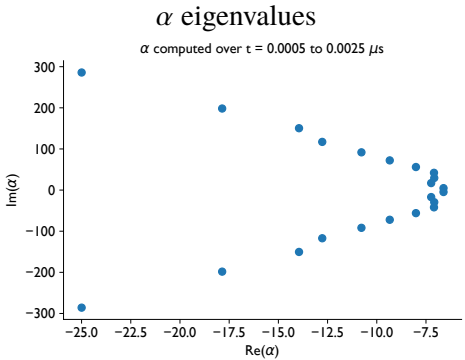


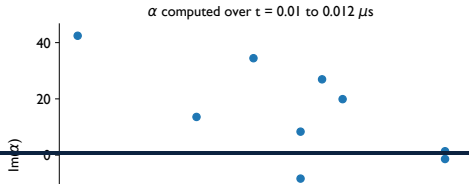


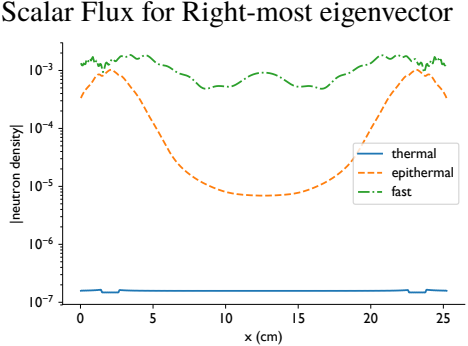
Left: Neutron density (ϕ/v) as a function of space and time.

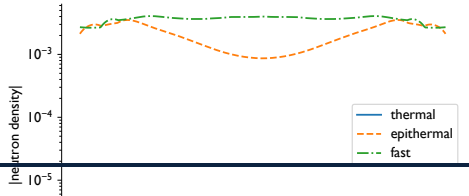
Right: Neutron spectrum at the center of the HDPE and the center of the fuel.

Using the DMD method at different times finds the predominant eigenmodes present in the system.

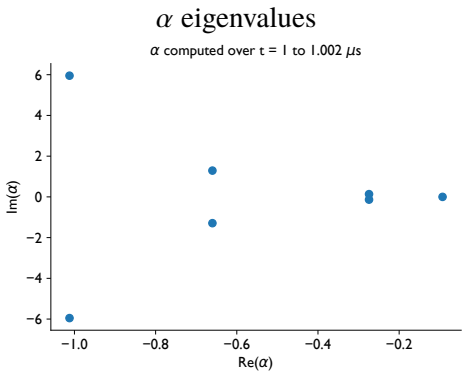


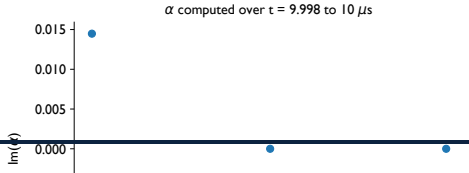


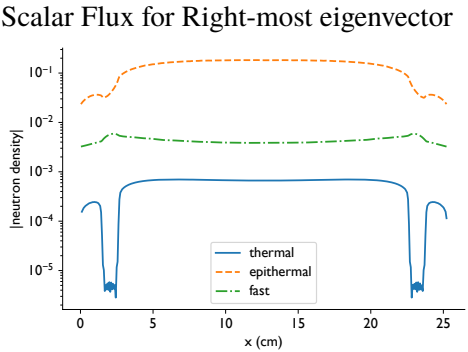


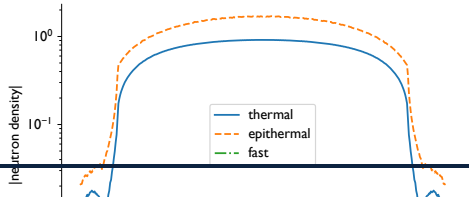


Using the DMD method at different times finds the predominant eigenmodes present in the system.



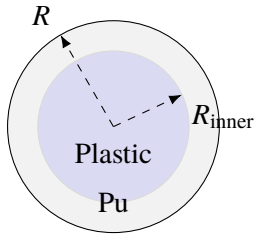


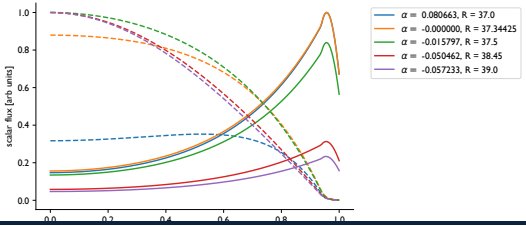


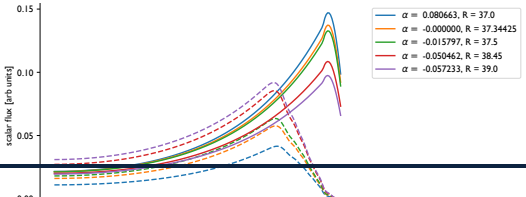


There is consistency between subcritical and supercritical DMD modes.

- We consider a two region sphere, where the mass is preserved but the outer radius is varied.
- The rightmost α eigenvalue changes character greatly when the criticality state is changed (left).
- The modes found by DMD in the first 0.1 μs of a simulation with initial condition of fast neutrons hitting the sphere are not as sensitive (right).







Extensions to the method could be made for more dynamic systems.

- Because the method finds the eigenmodes that are present in the system it will not be susceptible to finding unimportant, slowly decaying modes.
- One can tailor initial conditions to look for certain eigenvalues.
- The DMD approximation can be used for nonlinear operators.
- If the system changed over time we could approximate effective eigenvalues of the changing transport operator:
  - Thermal expansion
  - Depletion/Breeding
- The theoretical interpretation of these eigenvalues is not obvious.

Treating a linear problem using nonlinear methods can lead to improvement.

- The idea of using nonlinear methods for linear problems has a long history.
  - This can be traced back to at least Godunov's theorem that says that we need a nonlinear scheme to get a monotone, high order scheme for partial differential equations.
- Many problems can be cast in the form $\mathbf{y} = \mathbf{A}\mathbf{x}$, where $\mathbf{A}$ is a square matrix.
- However, it is possible to transform this to a under-determined system that has the same solution

$$\tilde{\mathbf{y}} = \tilde{\mathbf{A}}\mathbf{x},$$

  with the solution made unique by enforcing a nonlinear constraint or penalty on $\mathbf{x}$.
- This is the idea behind compressed sensing: if $\mathbf{x}$ is a sparse in some basis, we can randomly measure it and reconstruct the signal using a nonlinear optimization problem.

UQ/SA can be treated using this approach by using an optimization framework.

- In sensitivity analysis we are interested in finding the sensitivity of a quantity of interest (QoI) to a set of parameters: $\frac{\partial y}{\partial x_i}$
- Using a Taylor series we can write the QoI with $\bar{\mathbf{x}}$ as the nominal value of the parameters as

$$y(\mathbf{x}) = y(\bar{\mathbf{x}}) + \sum_{i=1}^{p} \frac{\partial y}{\partial x_i}(\bar{x}_i - x_i).$$

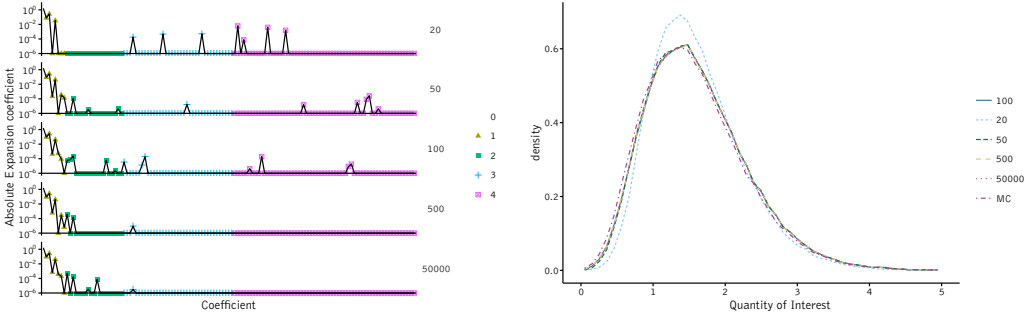
- We could compute the sensitivities, using finite differences (requiring $p + 1$ function evaluations), or we could compute $M < p$ random function evaluations $y(\mathbf{x})$ and solve the minimization problem

$$\min \sum_{i=1}^{p} \left| \frac{\partial y}{\partial x_i} \right| \quad \text{s.t. } \|\mathbf{A}\delta\mathbf{x} - \mathbf{y}\| < \epsilon,$$

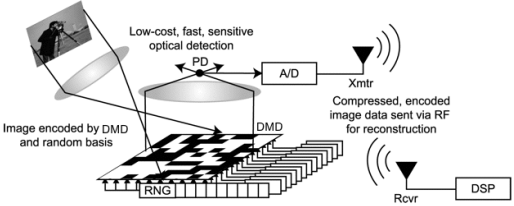
  with $\delta\mathbf{x}$ a vector of the sensitivities, $\mathbf{A}$ the $M \times p$ matrix containing the values of $(\bar{x}_i - x_i)$, and $\mathbf{y}$ the values of y for the M evaluations.
- This is equivalent to casting the problem of finding the sensitivities as a regression problem and solving it with regularized regression.
- We could extend this to give second-order sensitivities or the coeffs. of a polynomial chaos expansion.

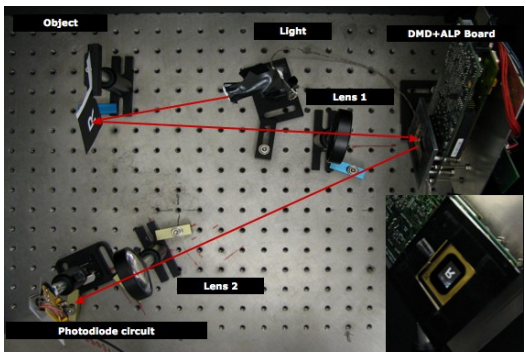
We can estimate a polynomial chaos expansion using many fewer function evaluations.

This example shows a fourth-degree polynomial expansion for a problem with two uncertain parameters.



- This is the idea behind the single-pixel camera developed at Rice University: sample the image projected onto random linear combinations of the pixels in the image using an array of tiny mirrors called a DMD. Each linear combination only requires the measurement of a single scalar value, i.e., a single CCD.





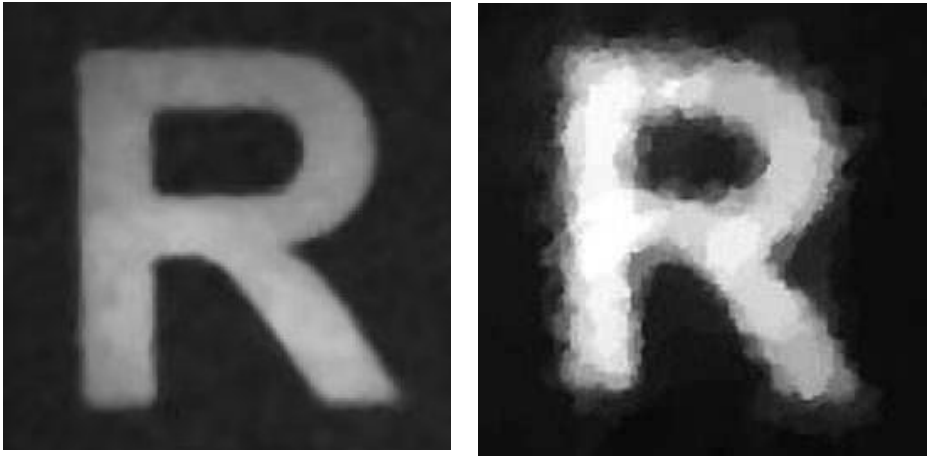
from http://dsp.rice.edu/cscamera

- In the case of the single pixel camera, each row of the measurement matrix $\mathbf{A}$ is a random string of zeros and ones of length N.
- The vector $\mathbf{b}$ is the sum of the pixels that are reflected to the detector by the mirror.
- In this case we find the solution to the system $\mathbf{Ax} = \mathbf{b}$ through the optimization problem

$$\text{minimize } \text{TV}(\mathbf{x}) \quad \text{subject to} \|\mathbf{Ax} - \mathbf{b}\|_2 \leq \epsilon,$$

  where the total variation (TV) of an image is the sum of the squares of the forward difference in the horizontal and vertical directions for each pixel (i.e., how much does the image change from pixel to pixel).
- This problem will have a unique solution, that ensures the reconstruction error (the second term) is small and the total variation in the image.
- Other optimization problems are possible, but for images, minimizing the TV norm is reasonable.

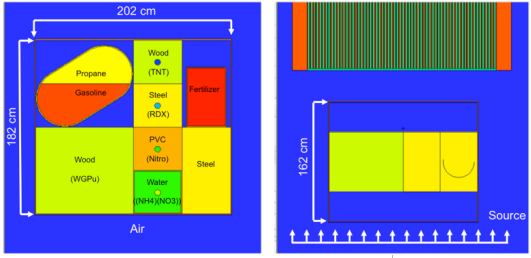
Single Pixel Camera Experimental Results

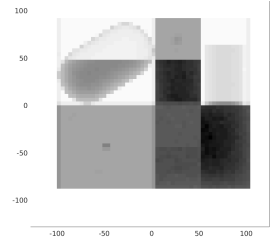


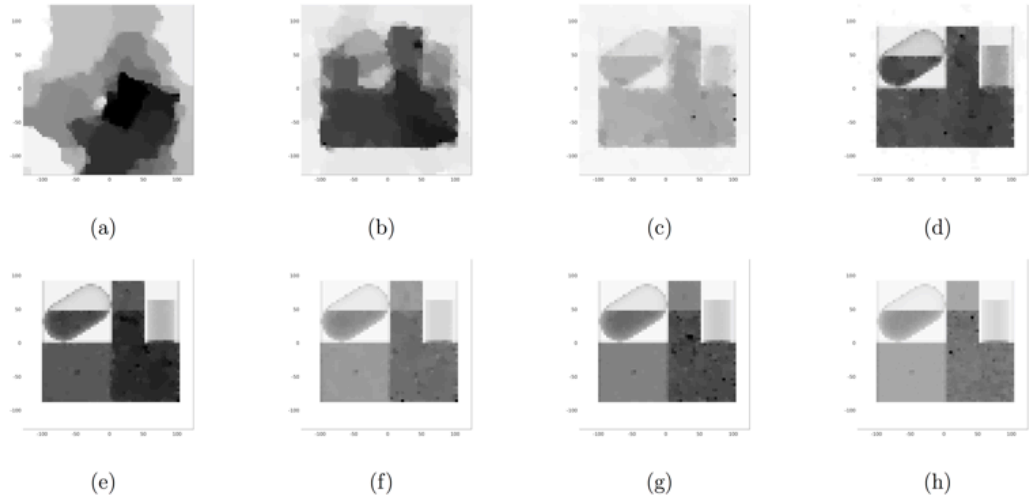
Left: Original 256 x 256 image, Right: Reconstruction from 1500 single-pixel samples (1/50)

- The actual matrix **A** does not need to be stored (just the random number seed used to generate it).
- The signal is encrypted because one needs to know **b** and **A** to determine the image.
- Trades computation for memory: the **b** vector is much smaller that the full image, but we must solve an optimization problem to recover the image.

- Consider a single, large neutron detector. We can fit this detector with a collimator and randomly block channels in the collimator to mimic the single-pixel camera.
- We consider an active interrogation problem of a cargo container with 14.1 MeV neutrons.
- We have a collimator that is a 3 m square and has $64 \times 64 = 4,096$ openings.



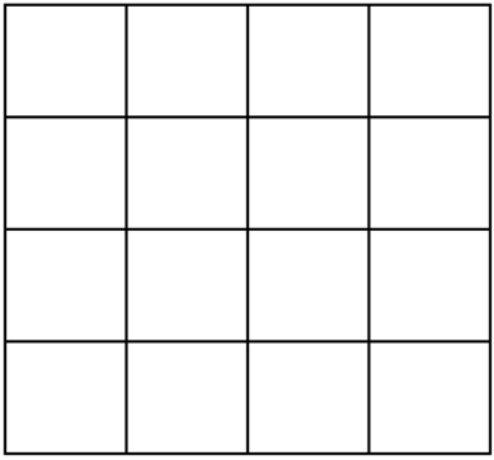




The logarithmic view of the reconstructions for the active interrogation problem using (a) 1%, (b) 5%, (c) 10%, (d) 20%, (e) 30%, (f) 40%, (g) 50%, and (h) 70% of the pixel count

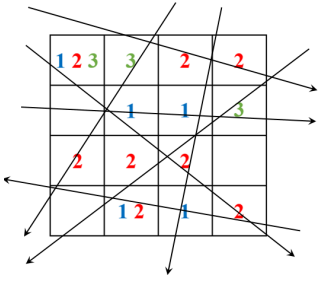
- One of the benefits of using the Monte Carlo method for particle transport simulations is that it uses simulated particles as analogs of the real particles in a system.
- As such we can create an analog of the single pixel camera inside of a Monte Carlo simulation to
  - Reduce the memory required to tally quantities, and
  - Reduce statistical noise in the simulation.
- The idea is to store random linear combinations of quantities estimated via Monte Carlo rather than a single value for each spatial location.
- These ideas could also be applied to MOC.

- If we want the solution on the grid on the right, we need a tally for each voxel.
- Instead, we could define a series of disjoint tallies that take linear combinations of the tallies for each voxel.
- If we need fewer disjoint tallies than the number of voxels, we can reduce the memory footprint for the calculation.

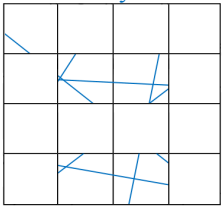


- If we want the solution on the grid on the right, we need a tally for each voxel.
- Instead, we could define a series of disjoint tallies that take linear combinations of the tallies for each voxel.
- If we need fewer disjoint tallies than the number of voxels, we can reduce the memory footprint for the calculation.

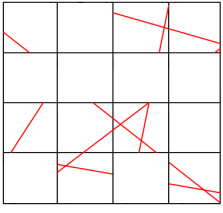
1 2 3	3	2	2
	1	1	3
2	2	2	
	1 2	1	2



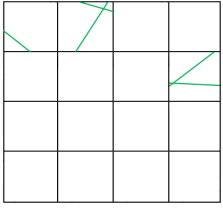
Tally 1



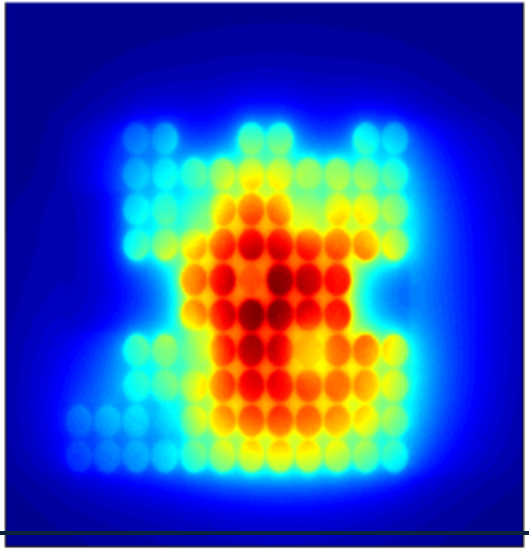
Tally 2



Tally 3

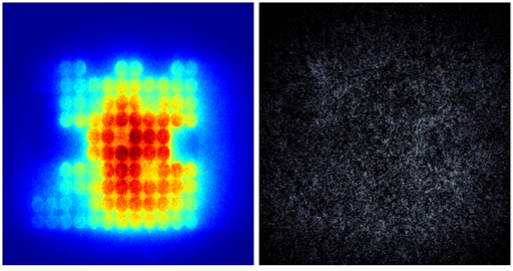


- As a test we look at a 2-D grid of the fast scalar flux in the Texas A&M TRIGA reactor.
- The grid has 1024×1024 voxels.
- Criticality calculation with 2,500 fission cycles and 200,000 neutrons per cycle.



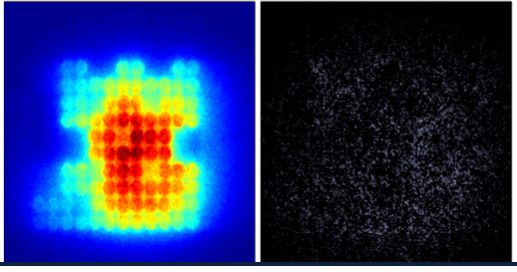
Reactor Calculation with Disjoint Tallies using 100x fewer particles shows reduction in memory/noise.

Original 2D mesh tally of flux in NSCR (left) and statistical error (right)



$\|e\|_1 = 0.0078$
$\|e\|_2 = 0.0014$
$\|e\|_\infty = 0.0071$

Compressed to 10% of computer memory (left) and statistical error (right)



$\|e\|_1 = 0.0069$
$\|e\|_2 = 0.0012$
$\|e\|_\infty = 0.0063$

Thinking about problems differently can lead to breakthroughs.

- In computation (especially transport) we often think that because we have the correct (often linear) model we should focus on solving that model better.
- However, using data to approximate that operator, or making things nonlinear can lead to improvement.
  - Approximate operators can tell us about what is happening in a given system.
  - Nonlinear reconstruction of tallies reduces memory and provides view-time computation.
  - Thinking of UQ and SA as a nonlinear problem may require fewer code runs.
- Nonlinear problems with less memory can be more efficient for advanced architectures.