SAND2021-10442PE

Sandia National Laboratories

Algorithms and Applications of the ITS Radiation Transport Code

ITS

CHEETAH MC

Brian Franke

with Kerry Bossler, Martin Crawford, Russell DePriest, Ron Kensek, Luke Kersting, Roger Martz, Aaron Olson

Computer Science Research Institute
Summer Seminar Series

August 26, 2021

U.S. DEPARTMENT OF ENERGY NNSA

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Outline

Monte Carlo radiation transport

ITS code capabilities

R&D topics

- Mesh geometry
- Sensitivities
- Stochastic media
- CHEETAH-MC (and Monte Carlo on GPUs)

Two totally different methods are available in computational physics to model radiation transport

Monte Carlo Methods (ITS)

Computer simulation of random walk by statistical sampling

- “Lagrangian” view: what happens to a given particle
- Runtime limited
  - Memory not generally a limitation
- Complex 3D modeling capability
- Efficient for computing integral quantities
  - Total charge crossing a surface
  - Total dose in a region
- Easily adaptable to traditional parallel computers (modern architectures are challenging)

Deterministic Methods (SCEPTRE)

Numerical solution of the mathematical equation describing the transport

- “Eulerian” view: what happens in a phase space element (r, E, Ω)
- Memory and/or runtime limited
- Complex 3D modeling capability
- Essential for computing differential quantities
  - Charge/energy deposition distributions
  - Space, energy, and angle dependent emission quantities
- Parallelizable, but challenging

Linear Boltzmann transport equation

The Monte Carlo method is sometimes said to solve the integral transport equation:

$$\psi(\vec{r},E,\vec{\Omega}) = \int T(\vec{r}' \to \vec{r},E,\vec{\Omega})S(\vec{r}',E,\vec{\Omega})dV' +$$
$$\int T(\vec{r}' \to \vec{r},E,\vec{\Omega}) \iint C(\vec{r},E' \to E,\vec{\Omega}' \to \vec{\Omega})\psi(\vec{r}',E',\vec{\Omega}')dE'd\vec{\Omega}'\,dV'$$

where S is a source, C is the collision operator, T is the transport operator:

$$T(\vec{r}' \to \vec{r},E,\vec{\Omega}) = \exp(-\Sigma_t(\vec{r}',E,\vec{\Omega})|\vec{r}-\vec{r}'|)\frac{\delta\left(\vec{\Omega}\cdot\frac{\vec{r}-\vec{r}'}{|\vec{r}-\vec{r}'|}-1\right)}{|\vec{r}-\vec{r}'|^2}$$

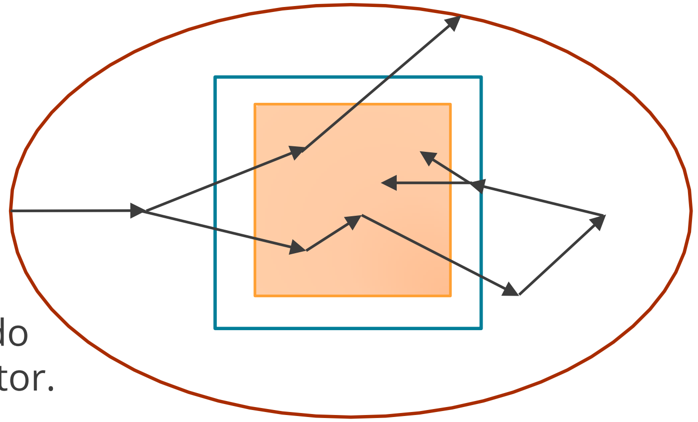
The Monte Carlo simulation can tally for the expected value of particle flux, but it can also tally a response, such as dose to a volume

$$D_{V_i} = \int\int \Sigma_{dep}(\vec{r}',E')\int \psi(\vec{r}',E',\vec{\Omega}')d\vec{\Omega}'dE'dV'.$$

One can also tally quantities that depend on correlations, such as the expectation that a particle history deposits total energy between E_j and E_{j+1} within volume V_i across all interactions. We call this a pseudo pulse-height detector.

The deterministic method solves the integro-differential transport equation:

$$[\vec{\Omega}\cdot\nabla + \sigma_t(\vec{r},E)]\psi(\vec{r},E,\vec{\Omega}) =$$
$$\int dE' \int d\vec{\Omega}'\sigma_s(\vec{r},E' \to E,\vec{\Omega}' \to \vec{\Omega})\psi(\vec{r},E',\vec{\Omega}') + q(\vec{r},E,\vec{\Omega})$$



Cartoon of a pseudo pulse-height detector.

Blue dots represent energy deposition events outside of the detector.
Red dots represent energy deposition events inside the detector.

Boltzmann-Fokker-Planck Equation

The continuous-energy condensed history code allows for external electric and magnetic fields.
The multigroup version of ITS includes continuous slowing down and continuous scattering terms.
We have been adding capabilities for time-dependence (as time-variation of sources, aging of particles during transport, and time-binning of tallies).

Only for charged particles:
Lorentz Force

$$\vec{F} = q\left(\vec{E} + \vec{v} \times \vec{B}\right)$$

Time-Dependence; External E&M Fields

$$\left(\frac{1}{v}\frac{\partial\psi}{\partial t}\right) + \vec{\Omega}\bullet\vec{\nabla}\psi + \left(\frac{1}{\gamma^3 m v}\vec{F}\bullet\vec{\nabla}_v\psi\right) + \sigma_t\psi =$$

$$\int dE' \int_{4\pi} d\vec{\Omega}'\,\psi(\vec{r},E',\vec{\Omega}',t)\sigma_s(\vec{r},E'\rightarrow E,\vec{\Omega}'\rightarrow\vec{\Omega}) + Q$$

Continuous Slowing Down; Continuous Scattering

$$+\frac{\partial}{\partial E}(S\psi) + \frac{\alpha}{2}\left\{\frac{\partial}{\partial\mu}\left[(1-\mu^2)\frac{\partial\psi}{\partial\mu}\right] + \frac{1}{1-\mu^2}\frac{\partial^2\psi}{\partial\varphi^2}\right\}$$

(restricted) stopping power

$$S(\vec{r},E) = \int dE' \int_{4\pi} d\vec{\Omega}'\,(E-E')\sigma_{ss}(\vec{r},E\rightarrow E',\Omega\rightarrow\Omega')$$

(restricted) momentum transfer

$$\alpha(\vec{r},E,\vec{\Omega}) = \int dE' \int_{4\pi} d\vec{\Omega}'\,(1-\vec{\Omega}\bullet\vec{\Omega}')\sigma_{ss}(\vec{r},E\rightarrow E',\Omega\rightarrow\Omega')$$

Nucl. Sci. Eng. **124**, (1996) pp. 369-389

Monte Carlo Methods

Monte Carlo methods are a class of computational algorithms for simulating the behavior of various physical and mathematical systems.

They are distinguished from other simulation methods by being stochastic (i.e., probabilistic - using pseudorandom numbers) as opposed to deterministic algorithms.

A classic use is for the evaluation of definite integrals, particularly multidimensional integrals with complicated boundary conditions

Because of the repetition of algorithms and the large number of calculations involved, Monte Carlo is a method suited to calculation using a computer.

The techniques were known as “statistical sampling” methods

- “Re-named” Monte Carlo methods in reference to the famous casino in Monaco by Stanislaw Ulam
- Ulam credits playing solitaire and pondering probabilities of winning with the inspiration
- Discussed idea with John von Neumann and planned first calculations on the new electronic computer

Even further back (18th century), George Louis Leclerc (Comte de Buffon) proposed the problem that has become known as “Buffon’s needle”

- Estimation of π by dropping a needle on parallel lines

Basics of Monte Carlo Methods

The method requires:

Knowledge of process, events, or function being simulated. This could include the following:

- Probabilities of event occurrences and outcomes
  - Interaction cross sections in radiation transport
- Mathematical description of a process

A suitable source of independent and identically distributed random numbers uniform on (0,1).

The ability to construct probability functions

- Discrete
- Continuous or Distributed

Let x be a physical variable

- $p(x)$ is a frequency function
- $p(x)dx$ = probability of x between x and $x + dx$

Also, let ξ be a RN uniform on (0,1)

- $p(\xi)$ is a frequency function
- $p(\xi)d\xi$ = probability of ξ between ξ and $\xi + d\xi$

Relate x-space to ξ-space by requiring

- $p(x)dx = p(\xi)d\xi$

Relating physical variable *x* to random variable ξ

Since ξ is uniformly random, it requires that *p*(ξ) be constant.

$$\int_0^1 p(\xi)d\xi = 1$$

$$C\int_0^1 d\xi = 1 \quad \Rightarrow \quad C = 1$$

$$C = 1 \quad \Rightarrow \quad p(\xi) = 1$$

Thus, *p(x)dx* = (1)*d*ξ. Through integration, we get

$$\int_{x_1}^{x} p(x')dx' = \int_0^{\xi} d\xi'$$

$$P(x) = \int_{x_1}^{x} p(x')dx' = \xi$$

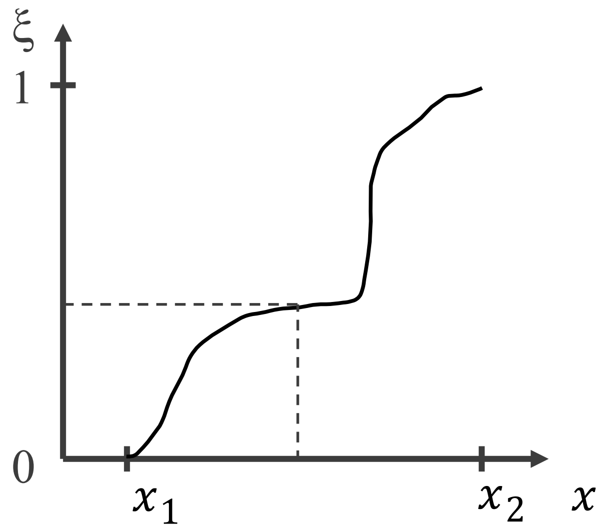
P(x) is a cumulative distribution function (CDF), varying monotonically from 0 to 1. It is the probability that *x'* is between x_1 and *x*. We want to compute the inverse:

$$x = P^{-1}(\xi)$$

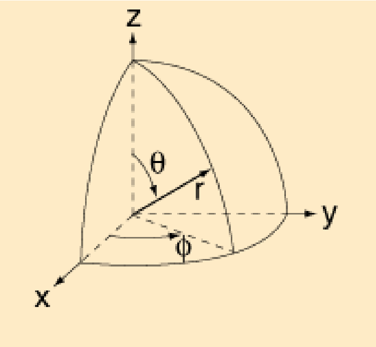
In some cases, the inversion can be done analytically, and the sample from the distribution is evaluated directly.

In some cases, techniques like rejection sampling can be devised to precisely sample from the correct distribution.

It is usually possible to sample from a numerical representation of the CDF.



Mathematics of Isotropic Scattering and Distance to Collision



First, we normalize to get a PDF.

$$\int_0^{2\pi} p(\varphi)\, d\varphi = 1 \quad \Rightarrow \quad p(\varphi) = \frac{1}{2\pi}$$

Then, we find the inverse of the CDF.

$$\frac{1}{2\pi}\int_0^{\varphi} d\varphi' = \int_0^{\xi_1} d\xi' \quad \Rightarrow \quad \varphi = 2\pi\xi_1$$

We use a similar approach for θ, but because we want to sample uniformly on the set of possible directions (surface area of a unit sphere), we must sample in cosine of theta ($\mu = \cos(\theta)$).

$$\frac{1}{2}\int_{-1}^{\mu} d\mu' = \int_0^{\xi_2} d\xi' \quad \Rightarrow \quad \mu = 2\xi_2 - 1$$

We use the property that the probability of a particle collision with a homogeneous background material is uniform and Markovian. As such, the probability that the next collision of a particle will occur between distance *s* and *s* + *ds* is:

$$p(s)ds = \Sigma e^{-\Sigma s} ds$$

We relate *s* and ξ_3 by the CDF

$$P(x) = \int_0^{\xi_3} d\xi' = \int_0^{x} \Sigma e^{-\Sigma x'} dx' \quad \Rightarrow \quad \xi_3 = 1 - e^{-\Sigma x}$$

And invert to obtain

$$x = -\frac{\ln(1 - \xi_3)}{\Sigma} \quad \Rightarrow \quad x = -\frac{\ln(\xi_3)}{\Sigma}$$

[Since ξ_3 is a uniform random variable on (0,1), so is (1 - ξ_3).]

Added Complexity

Physics

- Actual scattering distributions can be highly anisotropic for electrons and photons.
- Scattering distributions and secondary production distributions can have correlations in angle and energy.

Geometry

- Because transport is Markovian, a particle can be moved to a material boundary and the process reset.
- Particles must track to nearest of an interaction or geometry boundary.

Tallies

- The code must tally quantities of interest and provide statistical estimates.

Biasing

- Modifying the statistical sampling game can be done fairly to more efficiently provide the same expected value.
- Different techniques are needed for different physics and different problems.

Parallelization

- Monte Carlo for linear transport is “embarrassingly parallel” for traditional architectures.
- This is possible because we use domain replication. Memory usage can grow rapidly for mesh geometries or highly differential tallies.

ITS Particle Transport

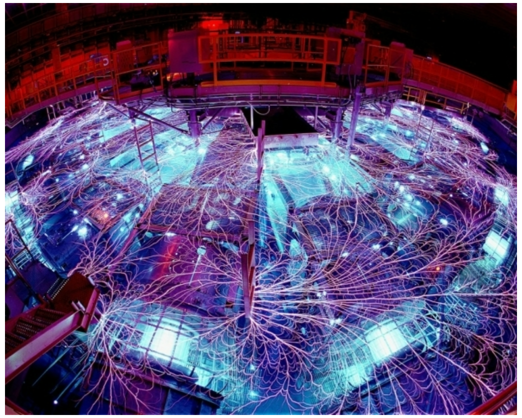
ITS primarily models high-energy photon, electron, and positron radiation. The same type of Monte Carlo approach is used for neutrons and ions.

In ITS, particle energies can range from 1 GeV to 1 keV (and, with many caveats, somewhat lower energies).

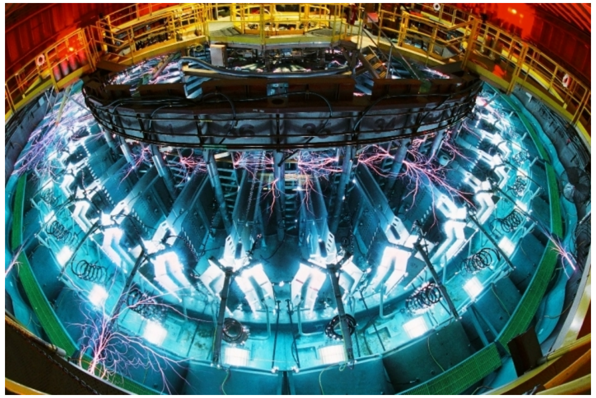
Particles interact only with unchanging background materials or fixed external fields. Particles do not interact with each other. It simulates linear Boltzmann transport.

It is used to simulate incident electron beams, gamma rays, x-rays, and the resulting electron-photon cascades.

Z pulsed-power machine

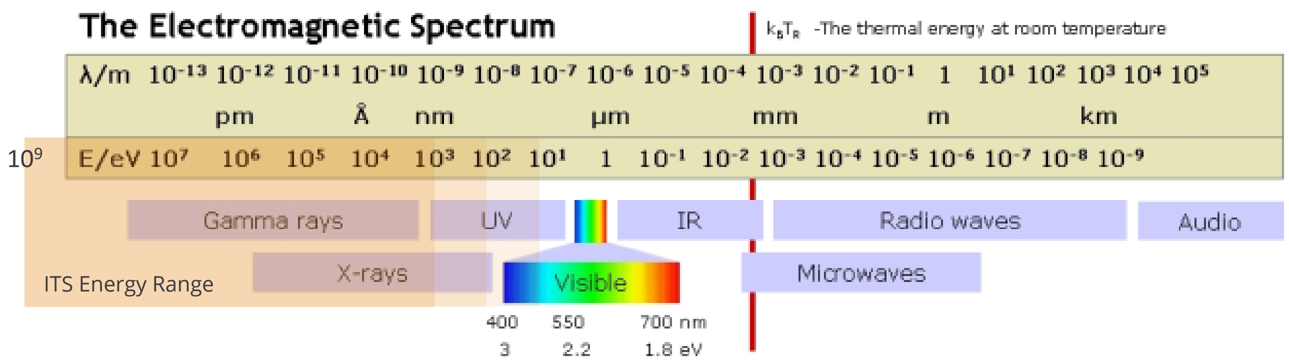


Saturn pulsed-power accelerator



High-Energy Radiation Megavolt Electron Source (HERMES) III accelerator

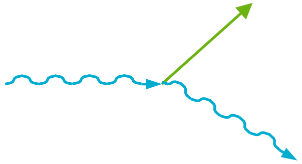
The Integrated TIGER Series (ITS) began in the early 1970s as the 1D (multimaterial!) TIGER code, as an extension of the NIST ETRAN code which was limited to a single material.
It supported the design of pulsed-power facilities and experiments at Sandia, but has been generalized for many other problems over time.



Photon Physics Models

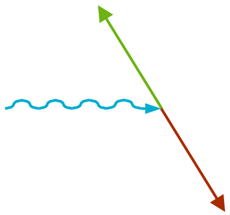
Photoelectric Absorption

- Produces photo-electron and atomic shell vacancy
- Vacancy may produce relaxation radiation



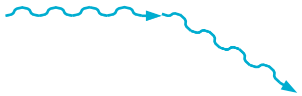
Incoherent Scattering

- Compton (with binding effects and Doppler broadening)
- Produces Compton electron, scattered photon, and atomic shell vacancy
- Vacancy may produce relaxation radiation



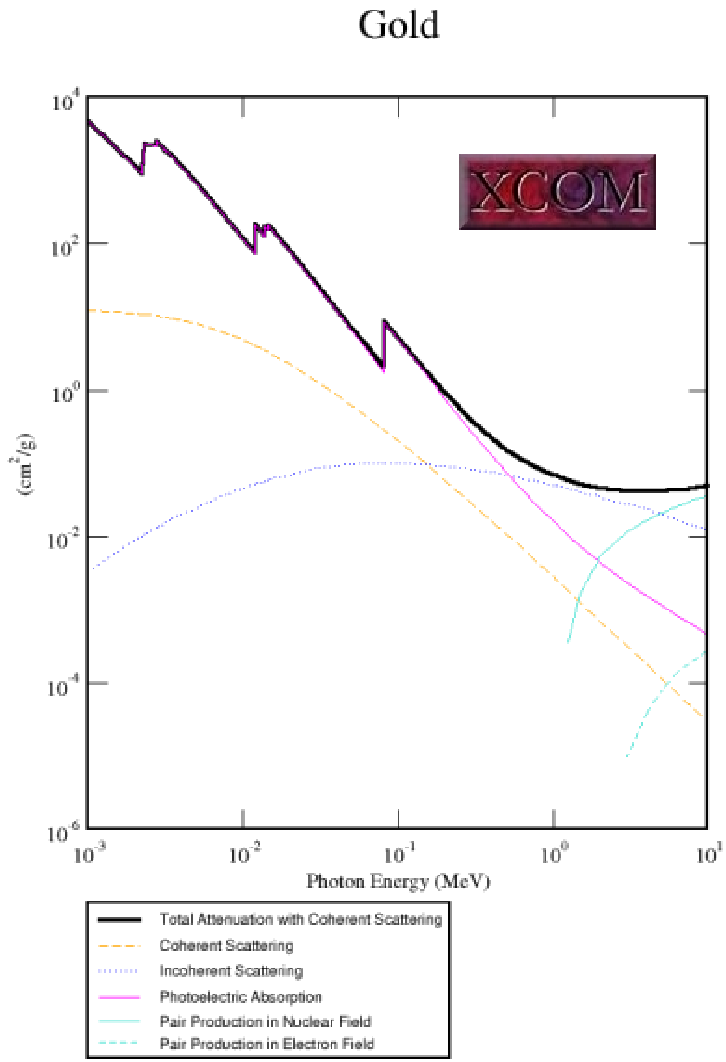
Pair Production

- Above the 1.022 MeV threshold
- Produces electron/positron pair

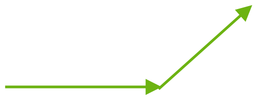


Coherent Scattering

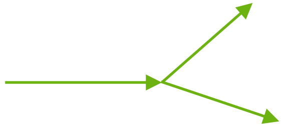
- Thomson scattering and binding effects
- Produces elastically scattered photon



Electron Physics Models

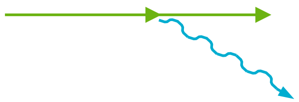


Elastic Scattering
- Produces deflected electron

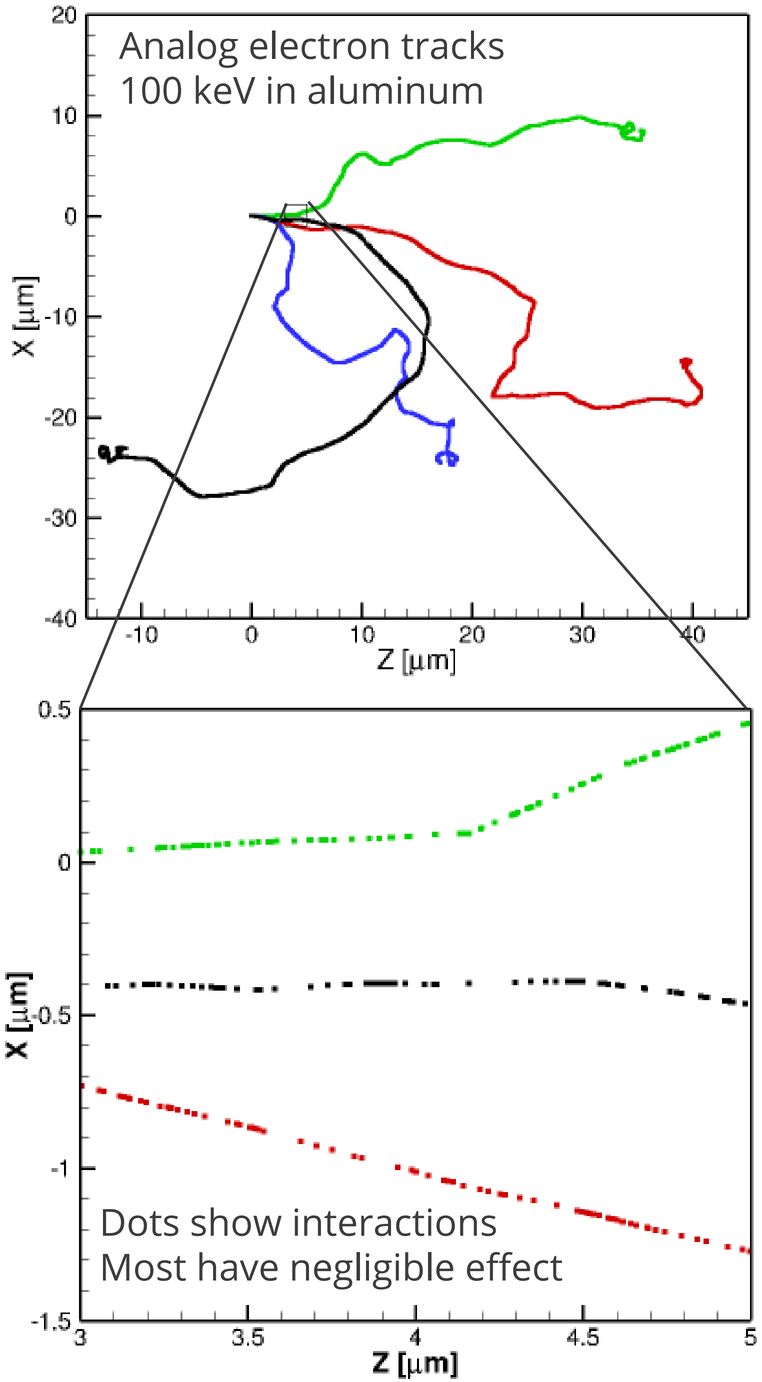


Inelastic Scattering
- Produces scattered electron, “knock-on” electron, and atomic shell vacancy
- Vacancy may produce relaxation radiation

Electronic Excitation
- Electron loses a small amount of energy
- Produces atom in excited state



Bremsstrahlung
- Electron “braking radiation”
- Produces photon

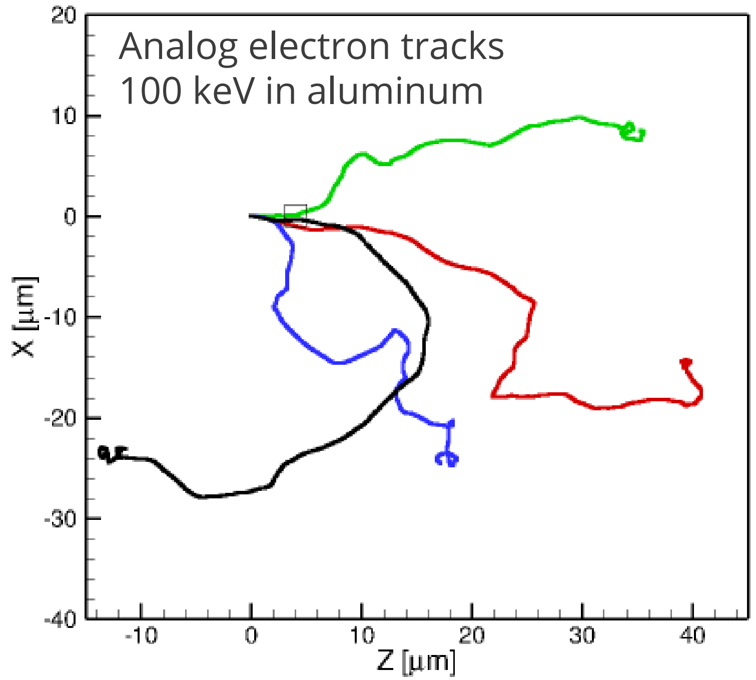


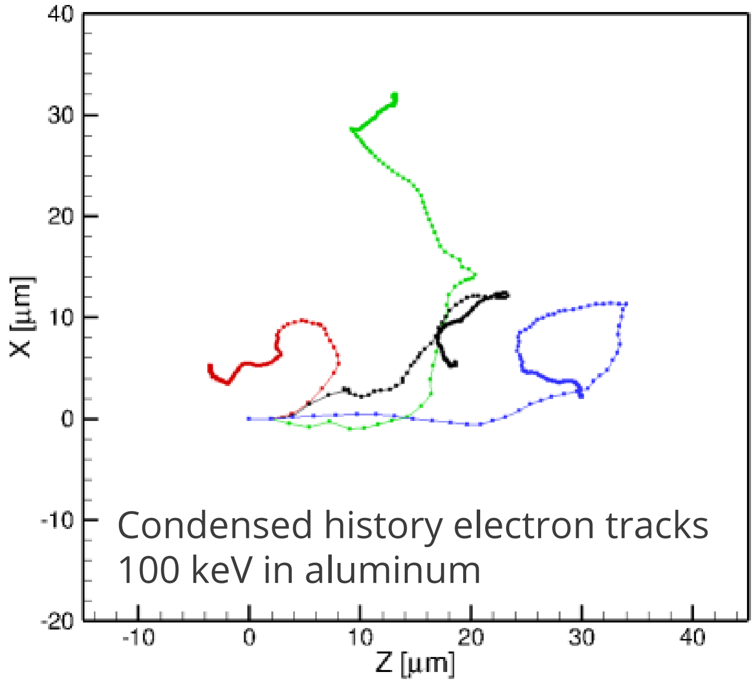
Condensed History Method

We can calculate the expected distribution of scattered particles after a predetermined pathlength, as the Goudsmit -Saunderson distribution.

- An approximation must be used to account for changes in energy, which changes the scattering distribution.
- The spatial displacement of the electron over that pathlength must be approximated.
- ITS uses the simplest spatial displacement approximation (move the electron directly forward and account for scattering at the end of the displacement).

A similar multiple-scattering approximation is used to determine the energy-loss of the electron for a specified pathlength.





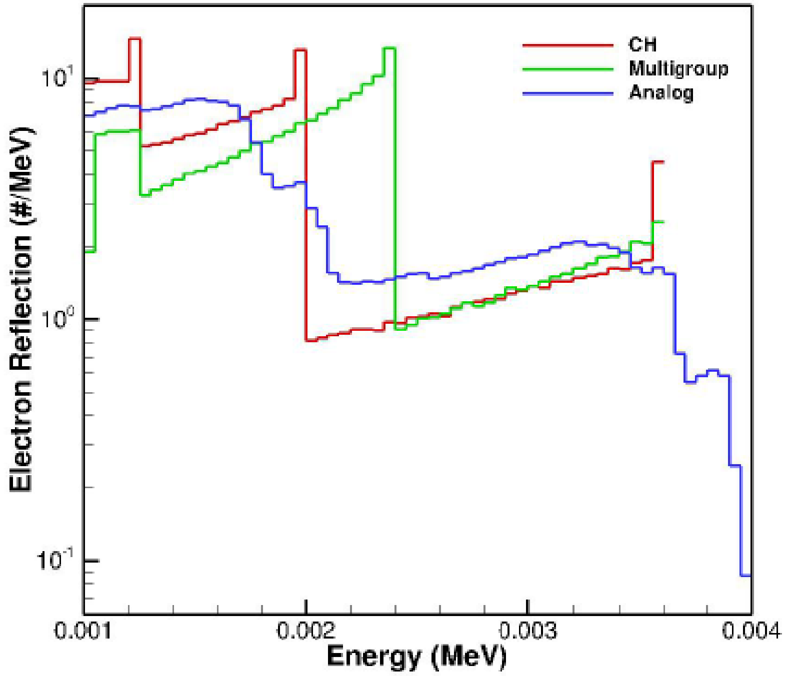
Forward and Adjoint; CH, Multigroup, Single-Scatter

We have three unique Monte Carlo capabilities.

Each has different strengths and weaknesses.

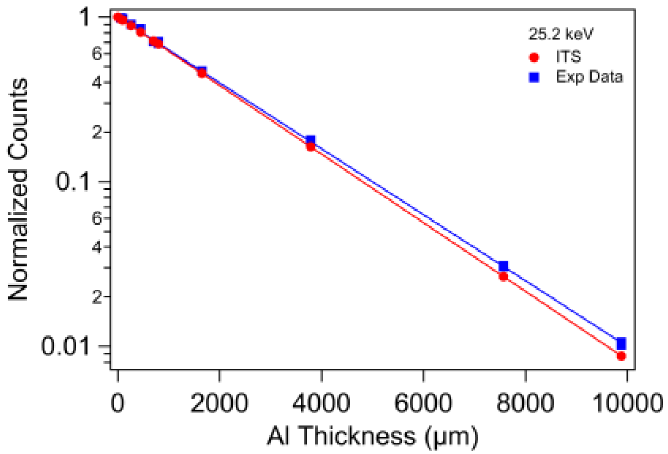
The single-scatter capability can be much more expensive, especially at high energies. We still have accuracy concerns with some of the cross section data. But the shell and relaxation data are much more detailed and may allow better predictions at low energies.

Differences in photo-electron reflection spectra from 4-keV photons on gold, with more detail included in the analog relaxation model.

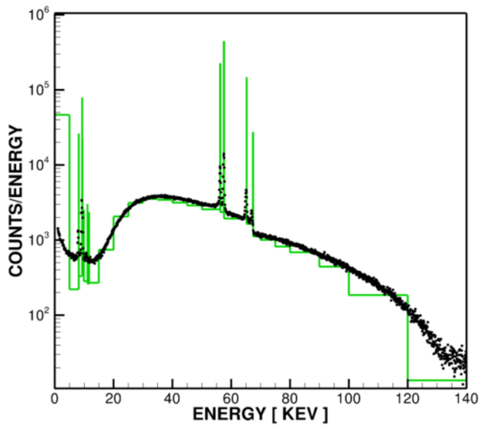


Sample ITS Validation

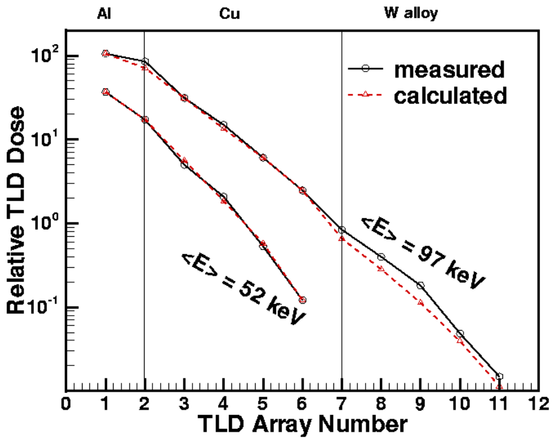
Photon Attenuation in Al



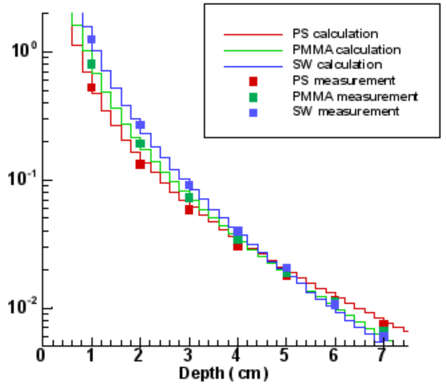
Simulating Linac Brems



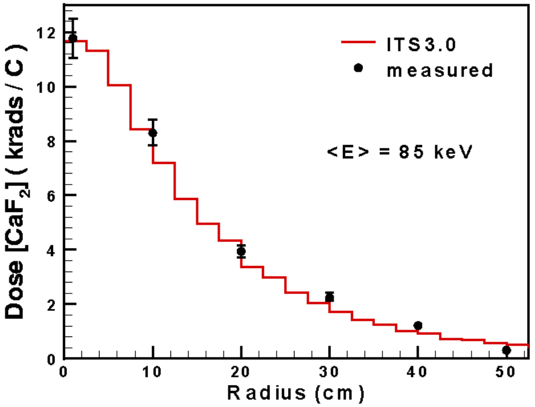
Brems Dose in Layers



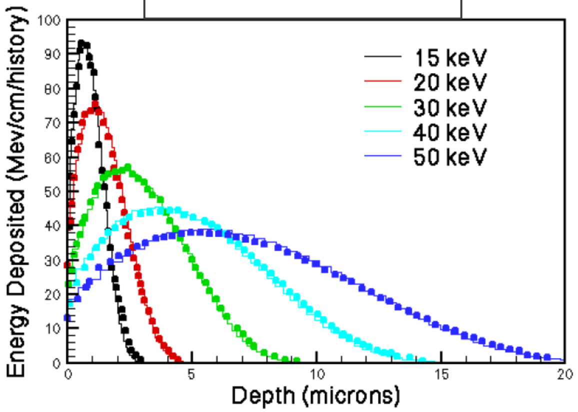
Composition Changes



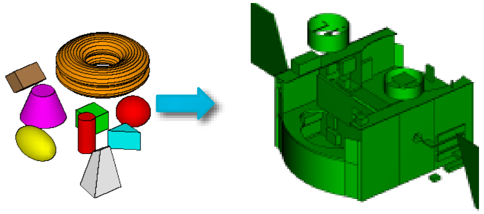
Brems Equilibrium Dose



Electrons in Si

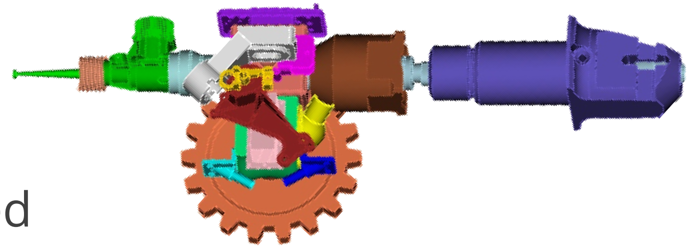


Geometry Capabilities



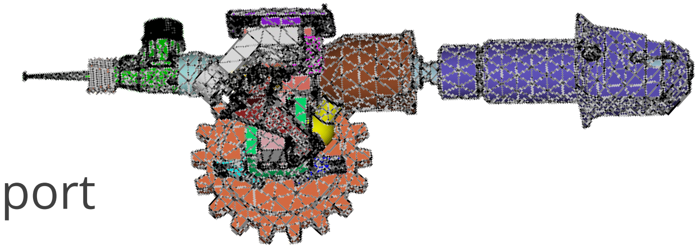
Native combinatorial solid geometry

- Boolean combinations of primitive bodies



CAD (ACIS® format)

- Can be as detailed as desired
- Separate ACIS® license is required
- ACIS® libraries not optimized for tracking: slower

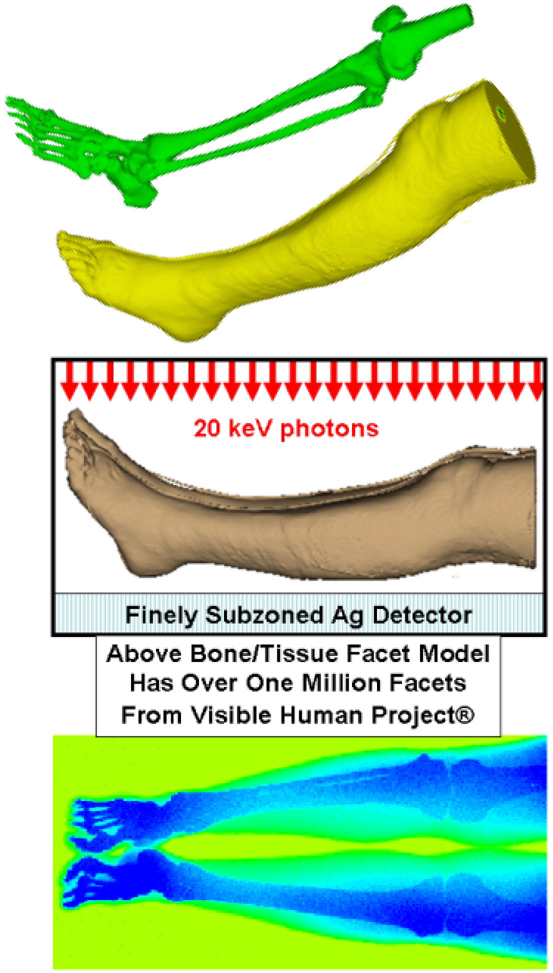


Facet-based geometry representation

- Cubit facet format
- Can use Cubit to surface mesh and export

Developing mesh capabilities…

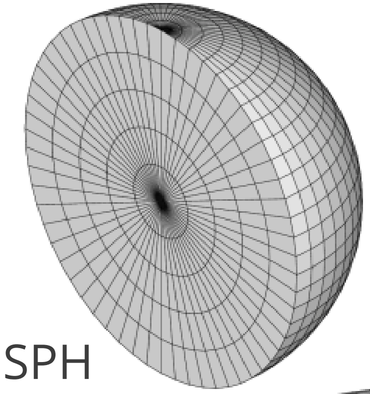
Combinations of the formats can be used in a single calculation.



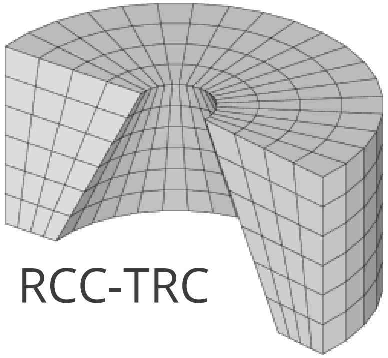
ITS-Simulated Dose(Ag) Distribution

3D Geometry "Subzoning": Tallies on Structured Meshes

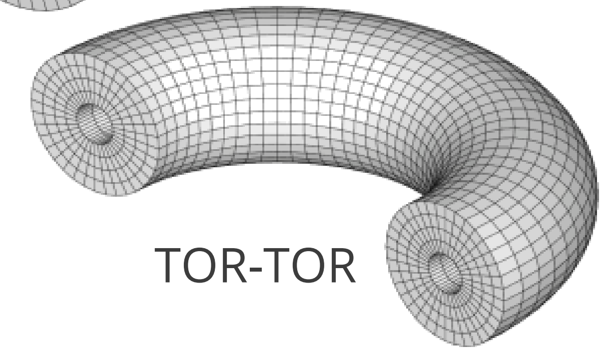
- Allows finer spatial resolution without having to cut up your geometry
- Structure saves on computation and memory
- Single-body and multi-body, conformal and non-conformal (overlay)
  - CAD body subzoning is always an overlay
- Similar "subsurfacing" capability for electron-emission tallies



SPH

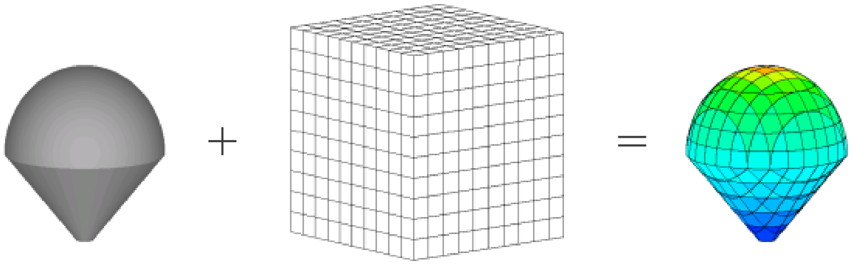


RCC-TRC



TOR-TOR

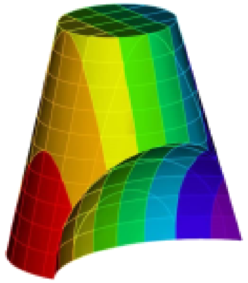
The reality of subzoning. This is what is happening internal to ITS.



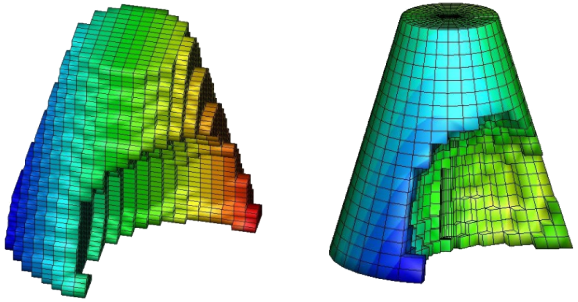
Particles are tracked on the zone geometry.

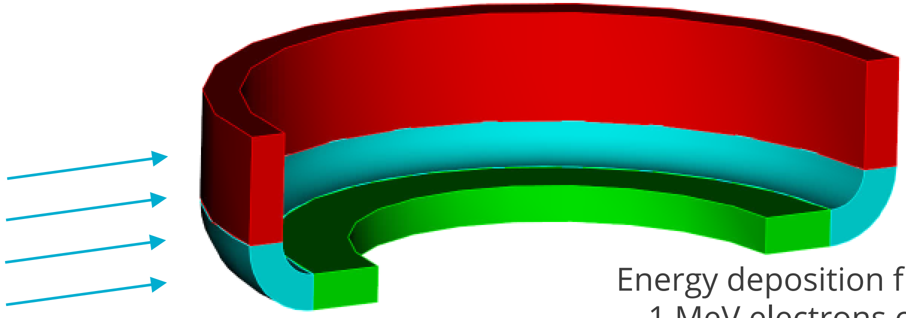
Zone tallies are made on the corresponding subzone structure.

Logical tally distribution

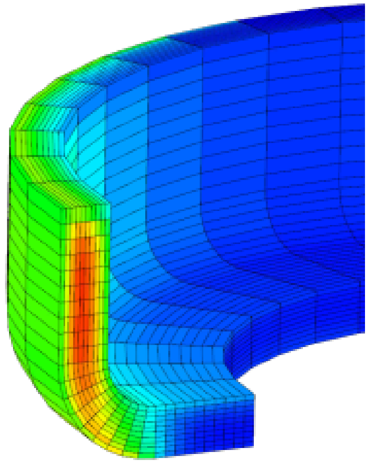


Mesh representation of tallies on non-conformal subzones





Energy deposition from 1 MeV electrons on cylindrical aluminum part



Adjoint Capability

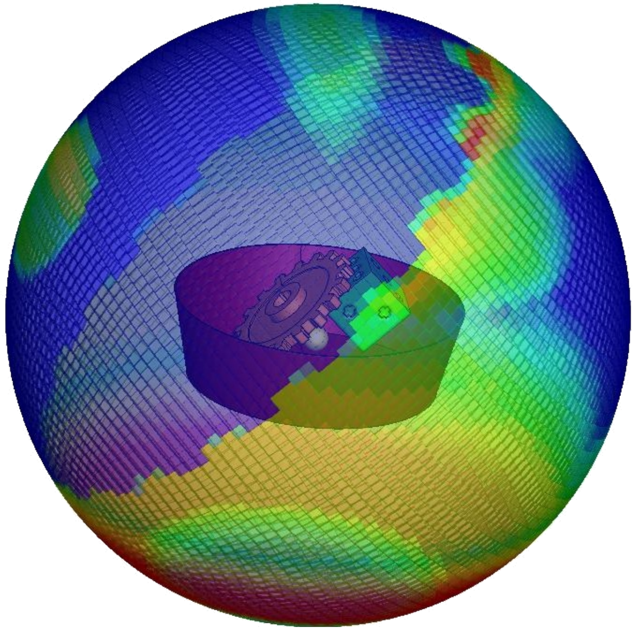
The 1D and 3D ITS codes have a multigroup capability that allows adjoint calculations.

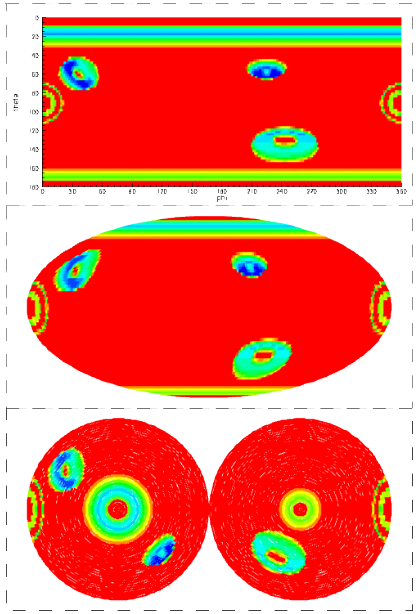
Adjoint advantages:

- Assessing a response from multiple sources with different space, energy, and angle distributions in a single calculation
- Generating response functions that can be used long after the initial calculation

There is an associated ray-tracing capability. (Images are all based on ray-trace results.)

- Facilitates mass-sectoring calculations
- Allows fast scoping of complex geometries





Plane Chart

Mollweide

Lambert Azimuthal

Adjoint Charge Deposition Validation

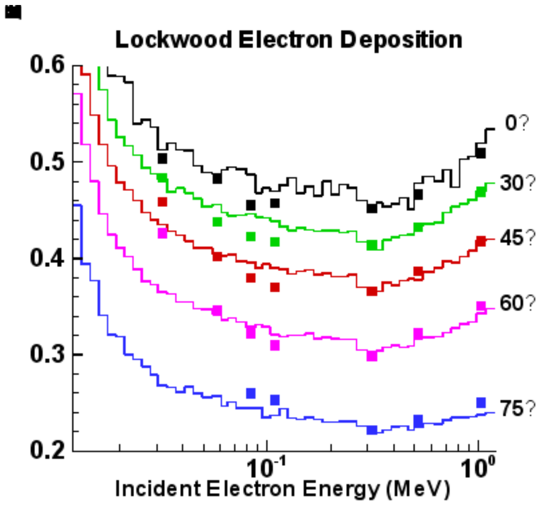
1-D Comparison With Experiments

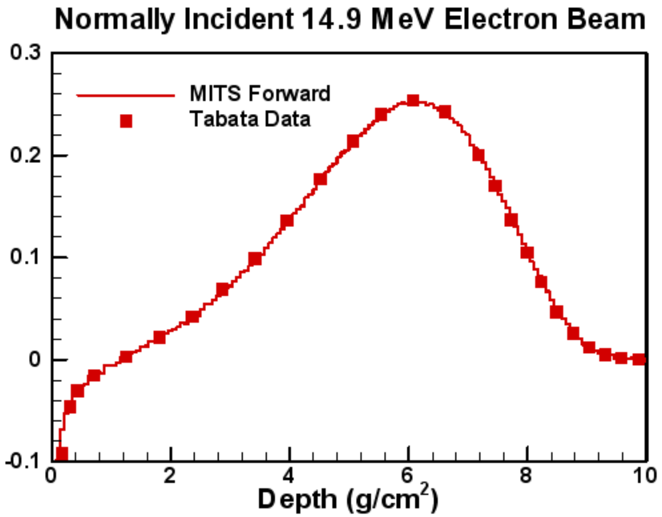
Volume Adjoint Charge Deposition

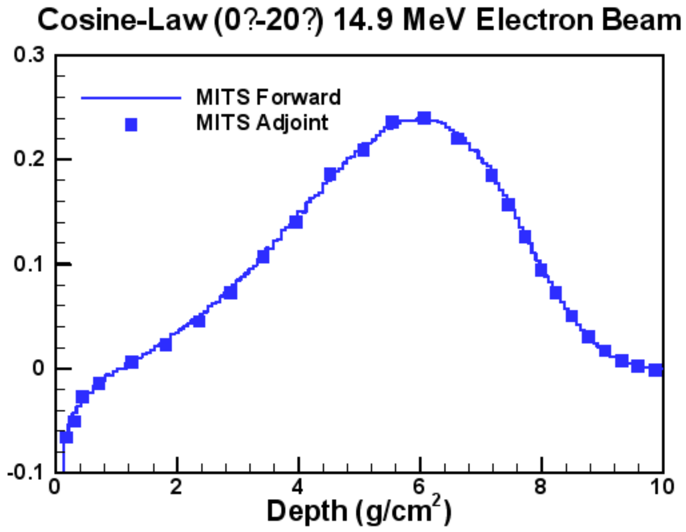
- Comparison with Lockwood Data

Point Charge Deposition

- Indirect comparison with Tabata Data







Pseudo Pulse-Height Simulations

ITS can calculate a quantity similar to the pulse -height quantity measured in proportional counters, in which the detector signal is proportional to the energy deposited by radiation.

ITS lacks some of the statistical variation of then electron-hole collection process in the detector, so we call it a pseudo pulse-height tally.

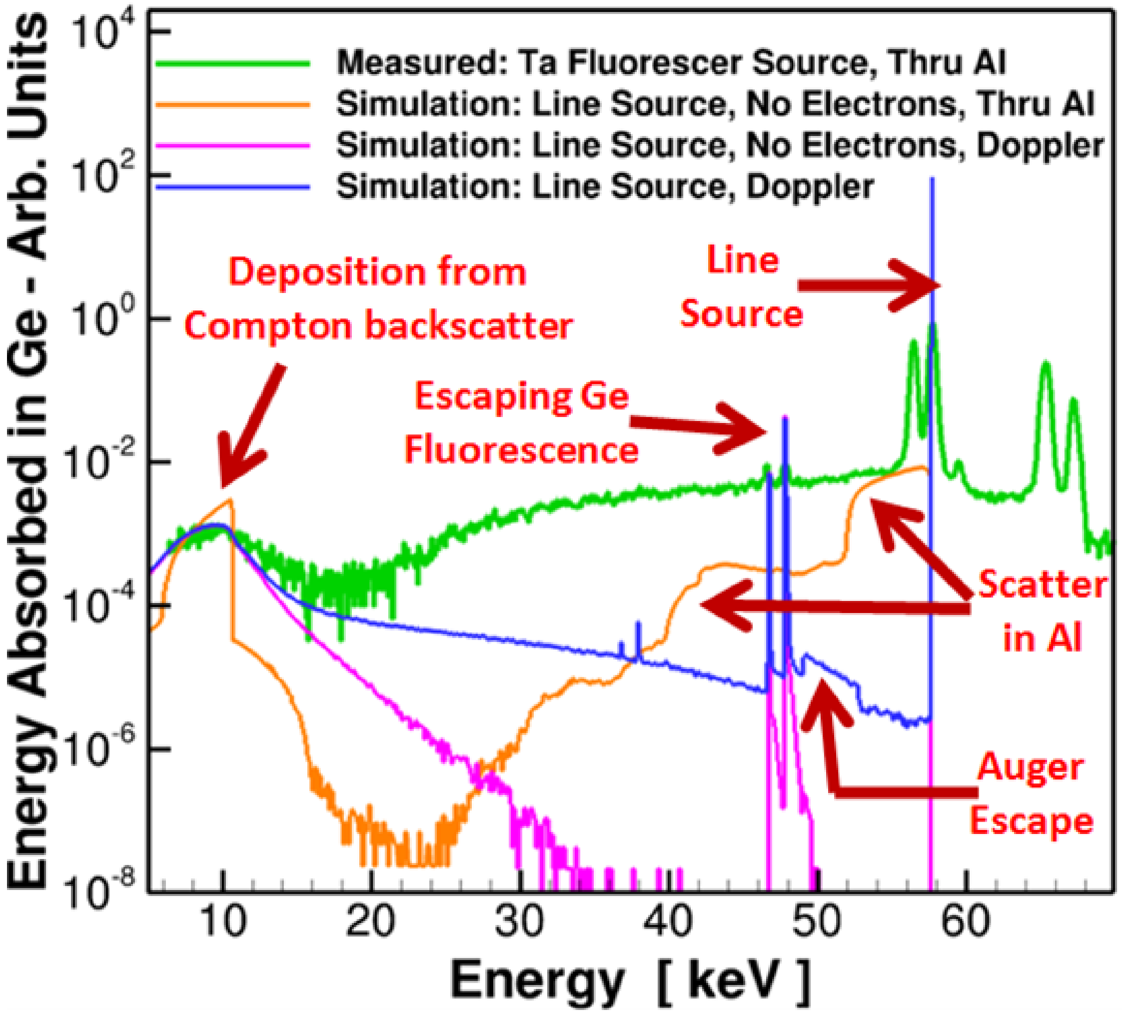


Figure: ITS simulations of a Ge pulse-height detector (spectrum of absorbed energy) due to a line source helped identify features of the measured spectrum (shown in red font).

Biasing

The goal of biasing is to maximize a statistical Figure-of-Merit:

$$\mathrm{FOM} = \frac{\bar{x}^2}{\sigma_{\bar{x}}^2 T}$$

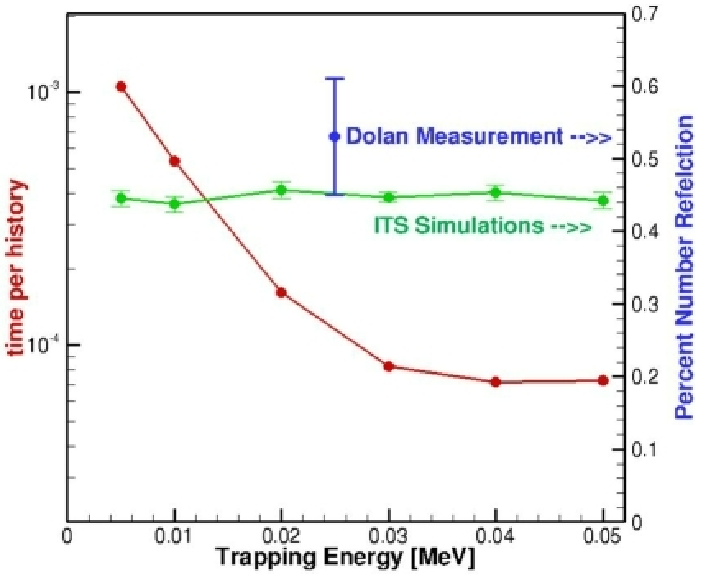
That is, minimization of the relative variance must be balanced with the minimization of the computational expense.

Another important objective is to minimize higher-order tally moments to minimize the occurrence of surprising outlier tallies (aka, "zingers").

While not technically biasing, in the sense of providing the same expected value, truncation methods often provide the greatest runtime savings with negligible effect on the expected value of tallies of interest.

Electron trapping is an automated feature in ITS for truncating low-energy electron transport that is often highly effective at increasing efficiency with negligible effect on accuracy.

ITS Variance Reduction Schemes	Type of Biasing
Electrons not tracked	Truncation
Electron trapping	Truncation
Cutoff energies	Truncation
Line radiation biasing	Truncation
Forced collisions	Modified sampling
Photon-produced electron roulette	Modified sampling
Bremsstrahlung production scaling (ITS)	Modified sampling
Impact ionization scaling (ITS)	Modified sampling
Photon-produced electron scaling (MITS)	Modified sampling
Electron-produced photon scaling (MITS)	Modified sampling
Source biasing mechanisms	Modified sampling
Next-event escape	Deterministic
Weight windows	Population control



Parallel Performance

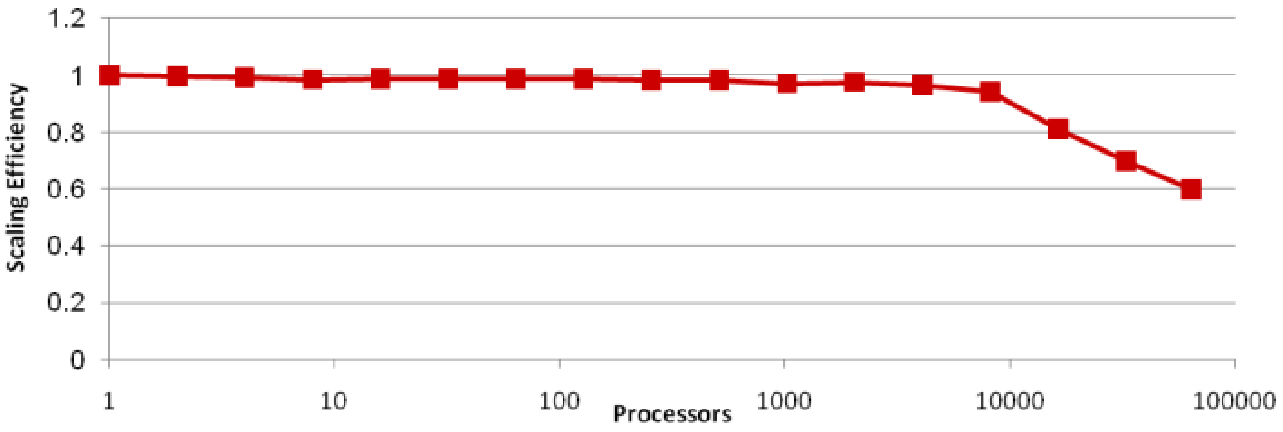
Scaling study is shown for a CAD-geometry photon-transport calculation performed on Cielo (LANL ASC machine) in about 2011.

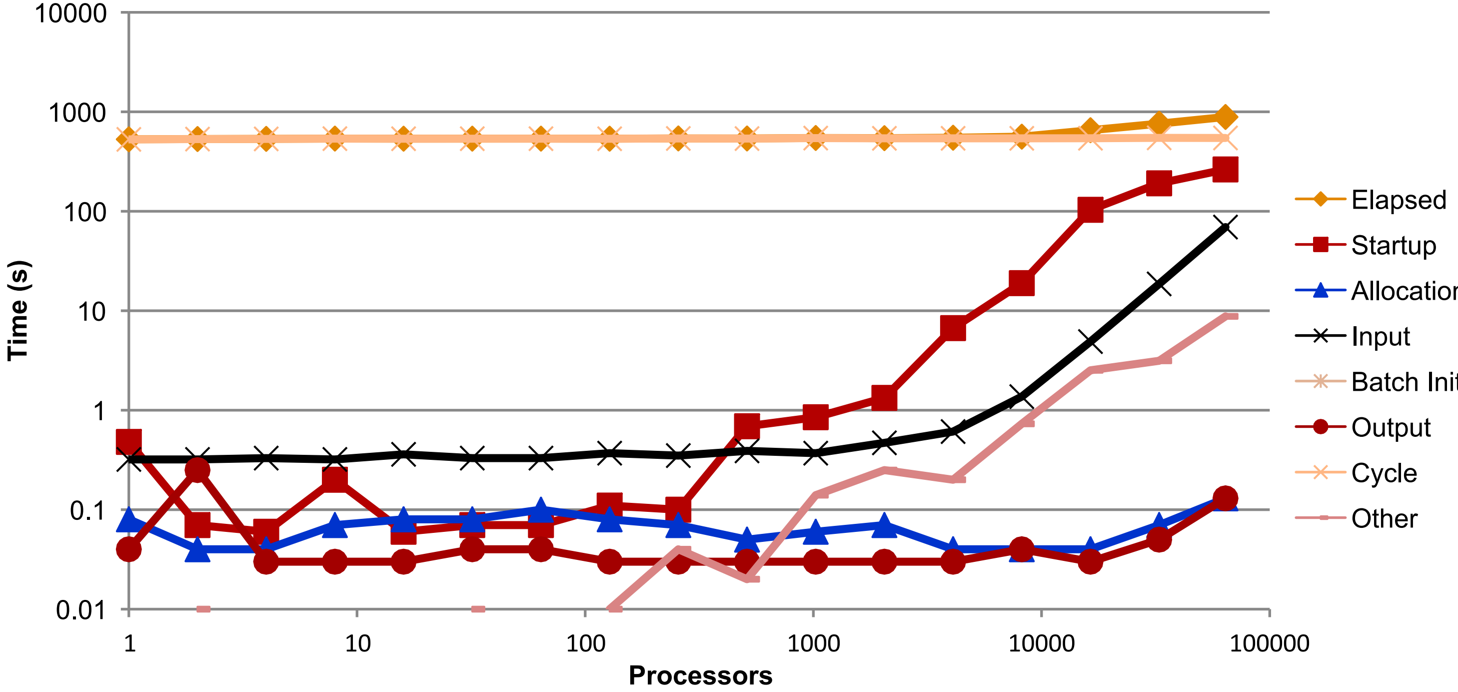
Loss of efficiency is primarily due to increasing MPI communication costs relative to the computational work. (Note that ideal runtime was less than 9 minutes.)

A sample of longer-running production calculations showed efficient but varying machine utilization:

94.4% (91.4 – 95.9) for static load balancing
96.9% for static load balancing with master idle
96.8% (93.7 – 98.6) for dynamic load balancing

Running on 2048 cores (with 16 cores per node) for about 20 hours with 20 restart checkpoints.





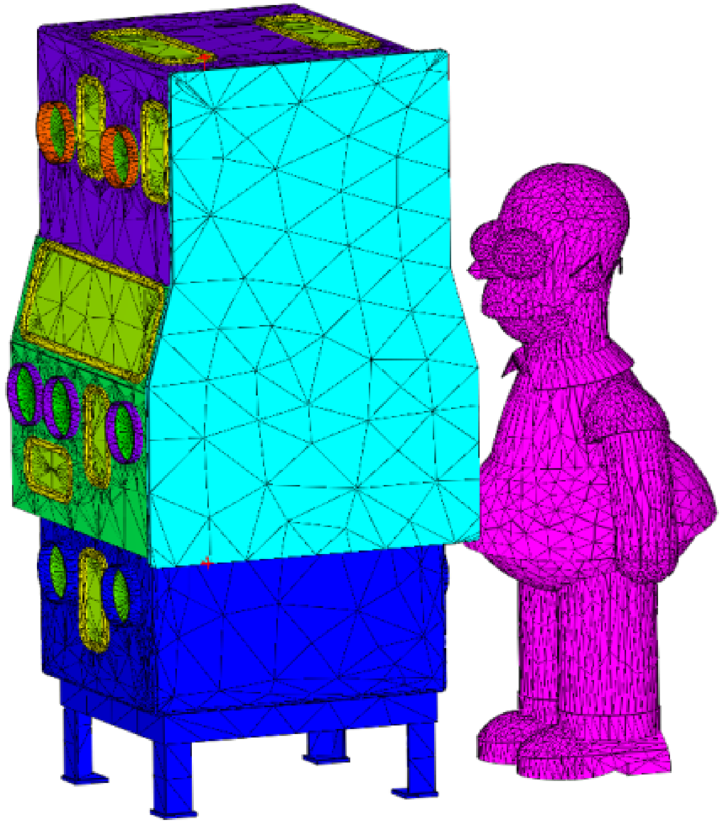
ITS Unstructured Mesh Tracking Library (UMTL)

Goal is to integrate seamlessly with combinatorial geometry (CG)

- Working with CAD geometry TBD

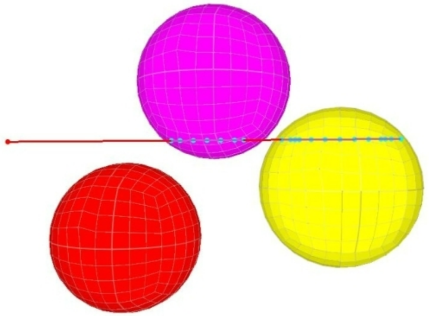
Tracking on 1st order tetrahedra, pentahedra, hexahedra

- Planning for 2nd order elements



Why Unstructured Mesh?

- Easily created with state-of-the-art 3-D CAD / CAE tools.
- Compared to CG, UM can more accurately model manufactured objects where surfaces were designed with splines, etc.
- Performance is slower than CG but faster than CAD
- Results tallied on the mesh are basically "free", avoid problems of non-conformal mesh overlays, and allow for state-of-the-art visualization.
- Easier "exchange of information" with mechanical engineering programs that use the finite element method based on unstructured meshes.



Test Problems:

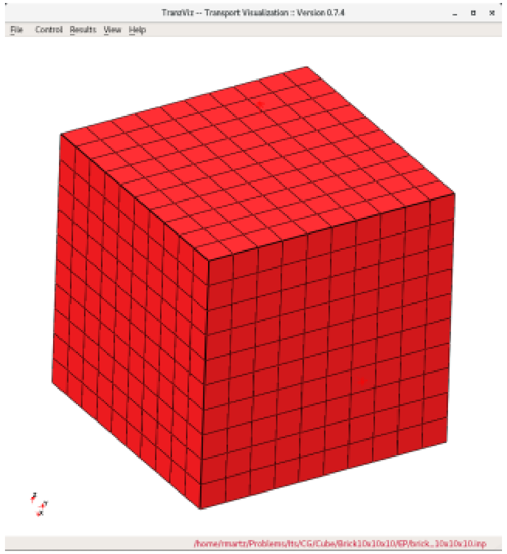
- Aluminum cube with varying number of sub-zones or hex mesh
- One-to-one correspondence sub-zone to hex element
- Exterior source impinging upon cube, with either electron source or photon source
- Photon performance is being investigated

Electron Source

Cube Detail	CG Time	UM Time	Ratio
8	16.51 (0.93%)	34.46 (1.16%)	2.09
1000	16.72 (1.39%)	38.80 (1.66%)	2.32
15625	16.90 (2.37%)	45.11 (0.84%)	2.67

Photon Source

Cube Detail	CG Time	UM Time	Ratio
8	12.32 (1.46%)	34.11 (0.70%)	2.77
1000	12.23 (1.85%)	115.50 (2.26%)	9.44
15625	11.99 (1.80%)	280.43 (1.26%)	23.40



Sensitivities

Under a REHEDS LDRD in 2016, we implemented and tested the differential operator technique, developed in the radiation transport community.

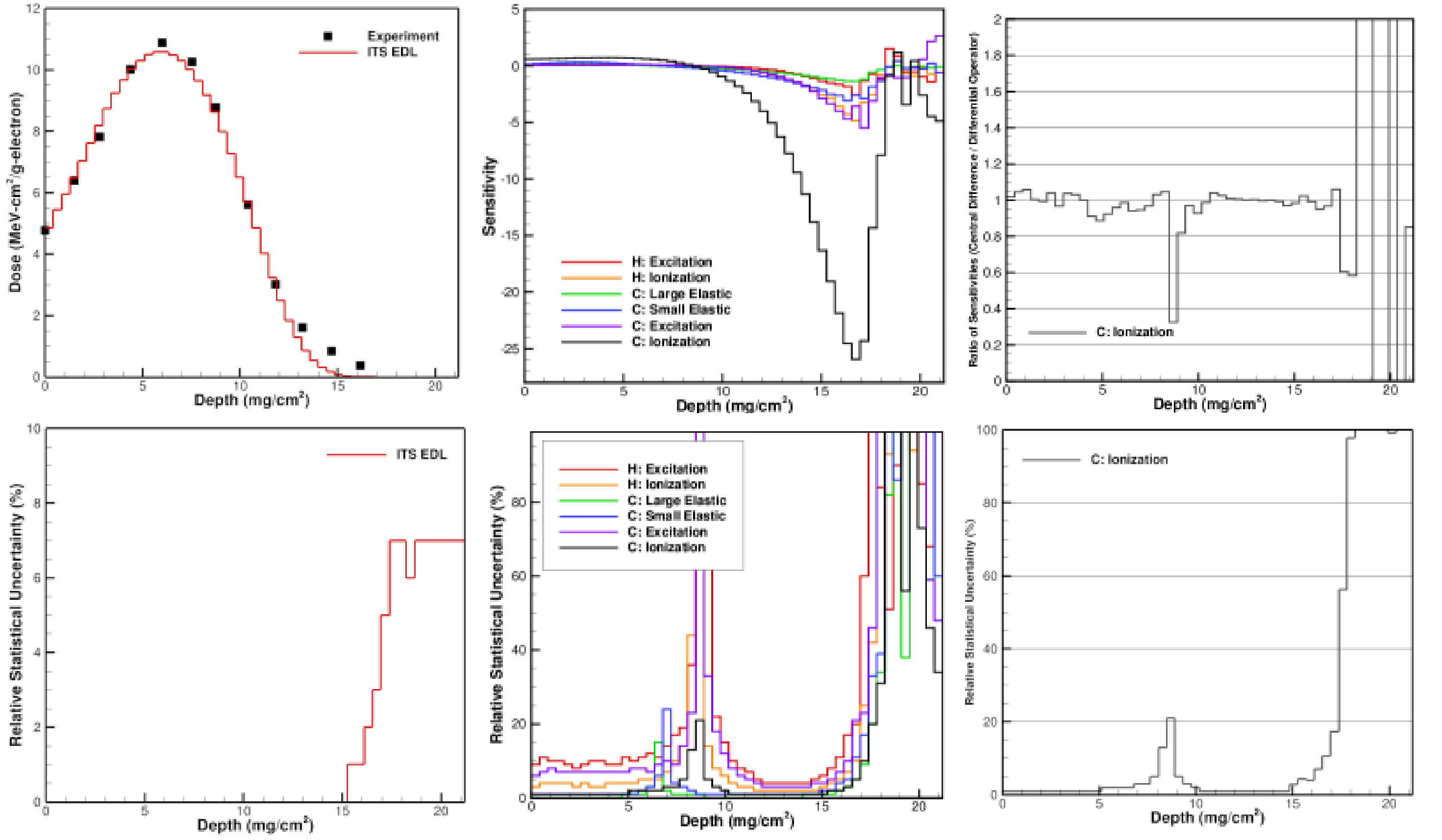
It can be effectively applied to calculating sensitivities with respect to some parameters, such as density, composition, and interaction cross sections.

It is specific to linear Boltzmann transport and does not apply to all problem parameters, such as geometry.

Under an ongoing CIS LDRD, we are investigating more general methods.

McLaughlin validation experiment of energy deposition from 100 keV electrons on polystyrene

Energy deposition vs. Experiment | Differential Operator Sensitivities | Differential Operator Sensitivities vs. Central Difference Sensitivities

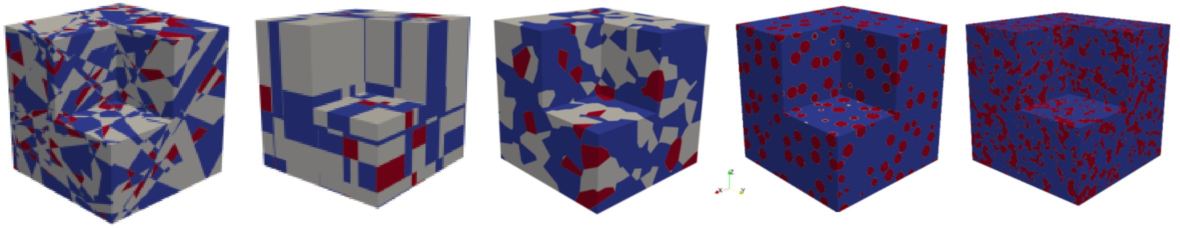


LABORATORY DIRECTED RESEARCH & DEVELOPMENT

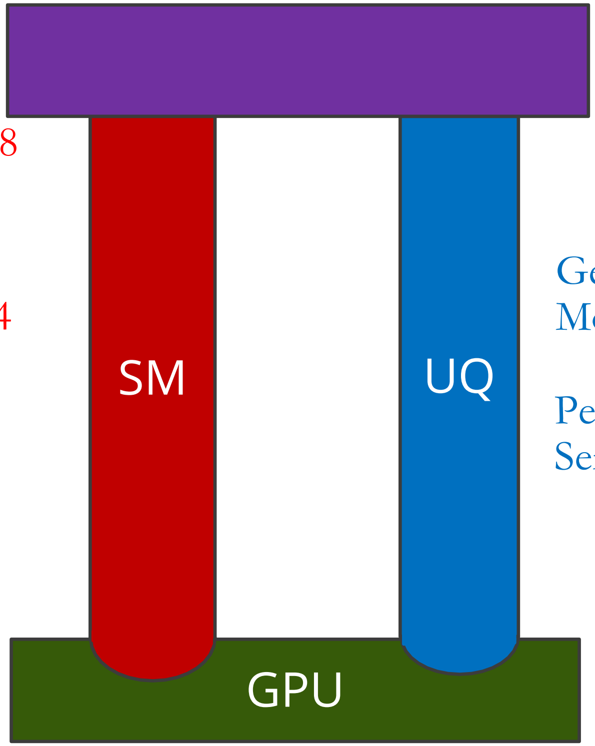
Overview: Next-Generation Monte Carlo Project

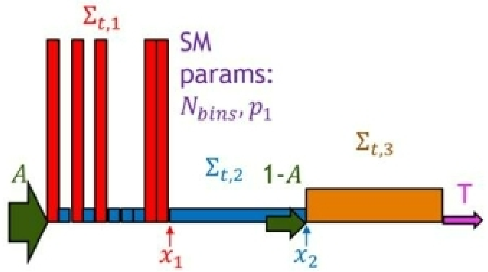
Develop efficient, embedded stochastic media (SM) and uncertainty quantification (UQ) Monte Carlo transport methods for the GPU.

Examples from five types of artificial stochastic media realization algorithms:



Poisson (Markovian) | Box-Poisson | Voronoi | Spherical Inclusions | Gaussian Process





Markovian three+ materials

Olson, Paper #33778
Transport in
Stochastic Media I

CoPS for generalized mixing

Davis, Paper #33784
Transport in
Stochastic Media II

CoPS memory/runtime efficiency

Vu, Paper #33614
Transport in
Stochastic Media II

Geraci, Paper #33671
Monte Carlo Methods

PCE surrogate models

Petticrew, Paper #33657
Sensitivity Analysis

Global sensitivity analysis

LDRD LABORATORY DIRECTED RESEARCH & DEVELOPMENT

CoPS on the GPU

Kersting, Paper #33673
Monte Carlo Algorithms

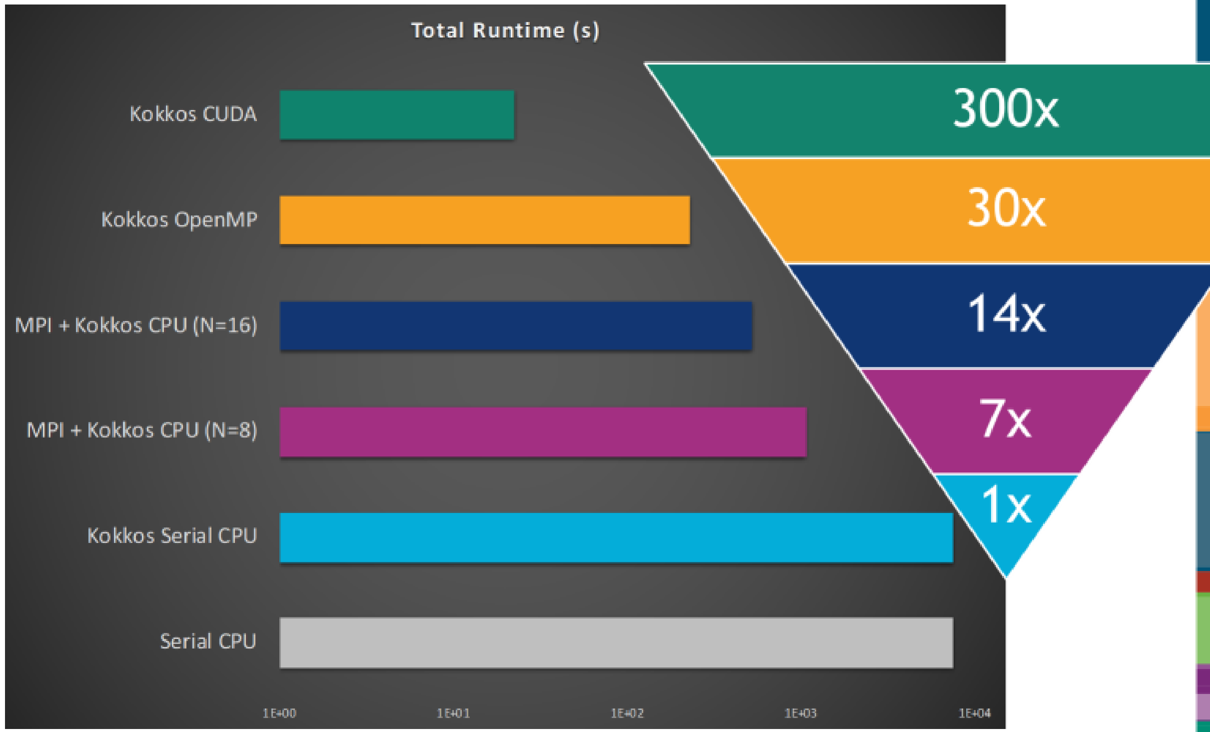
CHEETAH-MC

CHEETAH-MC is a new Monte Carlo code for photon and electron transport that can efficiently utilize both traditional CPUs and next generation platforms.

Feature	Available
Geometry	Combinatorial geometry Voxel geometry
Physics	Photon interactions Electron interactions Relaxation (FY22)
Particle Trackers	Woodcock tracking Surface tracking
Tallies	Event counters Conservation Energy and charge deposition Particle flux

Voxel geometry and Woodcock tracking are not ITS capabilities. They are expected to be useful for stochastic media.

Kokkos Performance for Simple Photon Absorption



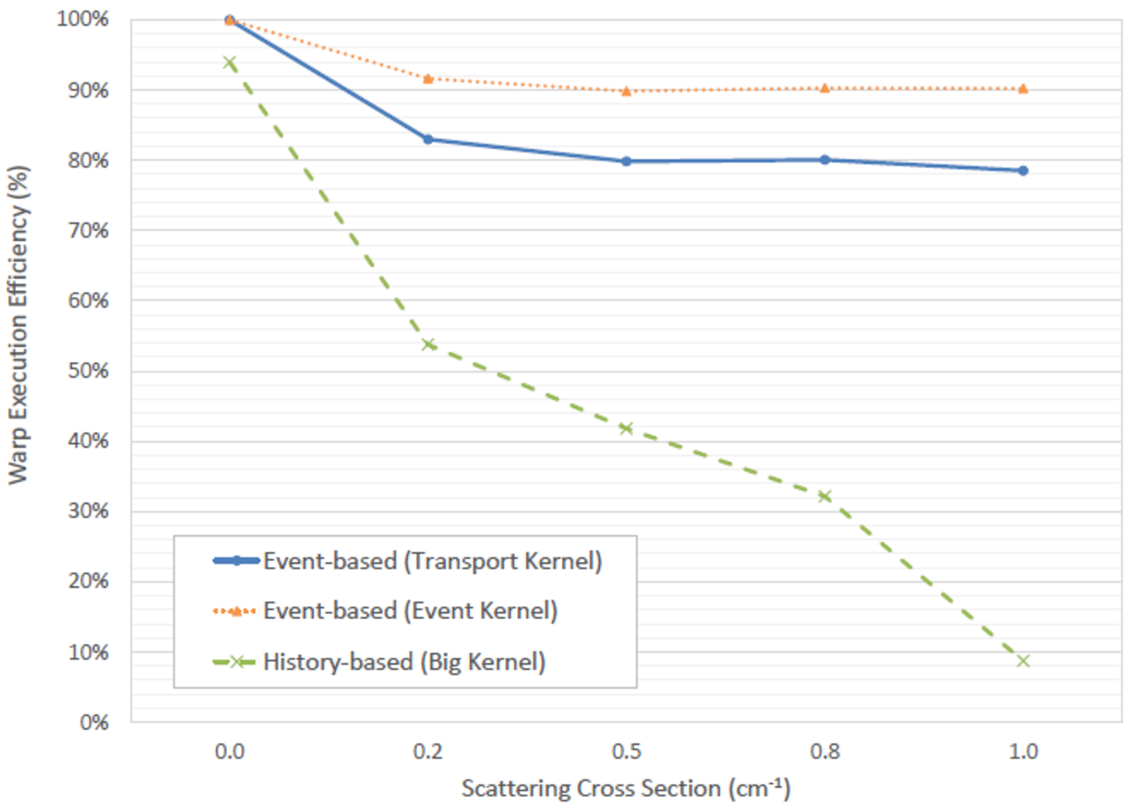
Simple photon absorption problem with 10^7 particle histories.

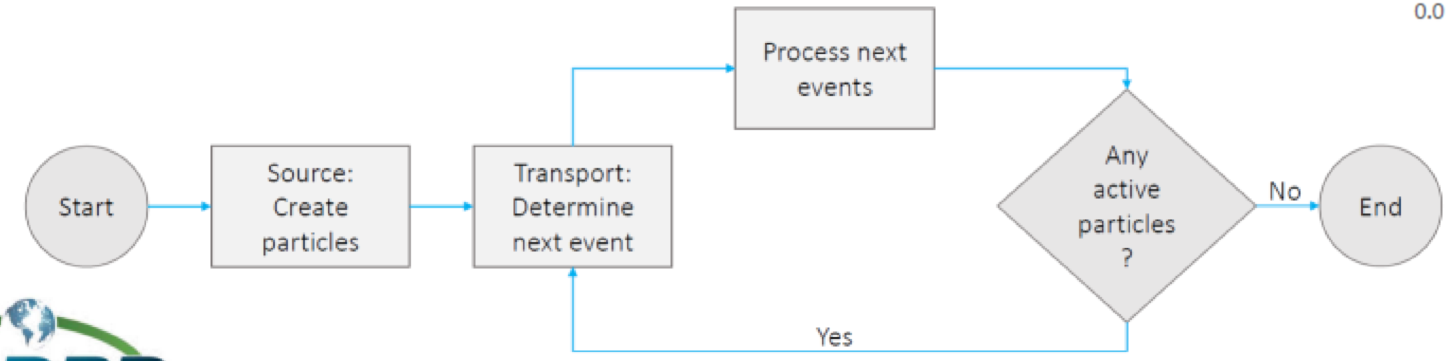
All cases were run on one P100 node on Ride and produced identical results to the serial CPU version.

No GPU or CPU optimization work has been done.

Reducing Divergence for Monte Carlo on GPUs

- Event-based Monte Carlo algorithm was explored in a research code
- Warp execution efficiency is a measure of branch divergence in the code
- Divergence for the Big Kernel increases dramatically as more scattering is added
- Transport and Event Kernel approach a fixed amount of divergence





Event-based algorithm significantly reduces divergence!

LABORATORY DIRECTED RESEARCH & DEVELOPMENT