Sunrise:

Panchromatic SED Models of Simulated Galaxies

Lecture 3:

Monte Carlo Radiation Transfer

Patrik Jonsson

Harvard-Smithsonian Center for Astrophysics

Lecture outline

- Lecture 1: Why Sunrise? What does it do? Example science. How to use the outputs? Projects?
- Lecture 2: Sunrise work flow. Parameters, convergence, other subtleties.
- Lecture 3: Radiation transfer theory. Monte Carlo. Polychromatic MC.
- Lecture 4: Dust emission, dust self-absorption. Sunrise on GPUs. Science.

The equation of radiative transfer

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \hat{\mathbf{k}} \cdot \nabla I_\nu = \frac{1}{4\pi}\rho j_\nu - \rho\kappa_\nu^{\mathrm{abs}} I_\nu - \rho\kappa_\nu^{\mathrm{sca}} I_\nu + \rho\kappa_\nu^{\mathrm{sca}} \int \phi_\nu(\hat{\mathbf{k}}, \hat{\mathbf{k}}') I_\nu(\hat{\mathbf{k}}') \mathrm{d}\Omega'$$

where

$$\mathrm{d}E = I_\nu(\hat{\mathbf{k}}, \mathbf{x}, t)\hat{\mathbf{k}} \cdot \mathbf{dA}\, \mathrm{d}\Omega\, \mathrm{d}\nu\, \mathrm{d}t$$

- Ouch... Simplify by ignoring time dependence and only looking at the intensity in a specific direction:

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}x} + \rho\kappa_\nu I_\nu = \rho\left(\frac{j_\nu}{4\pi} + \kappa_\nu^{\mathrm{sca}}\Phi_\nu\right)$$

from $$\frac{\mathrm{d}I_\nu}{\mathrm{d}x} + \rho\kappa_\nu I_\nu = \rho\left(\frac{j_\nu}{4\pi} + \kappa_\nu^{\mathrm{sca}}\Phi_\nu\right)$$

define the **optical depth**

$$\mathrm{d}\tau = \rho\kappa_\nu\,\mathrm{d}x$$

and we get

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau} + I_\nu = \frac{j_\nu}{4\pi\kappa_\nu} + \frac{\kappa_\nu^{\mathrm{sca}}}{\kappa_\nu}\Phi_\nu \equiv S_\nu$$

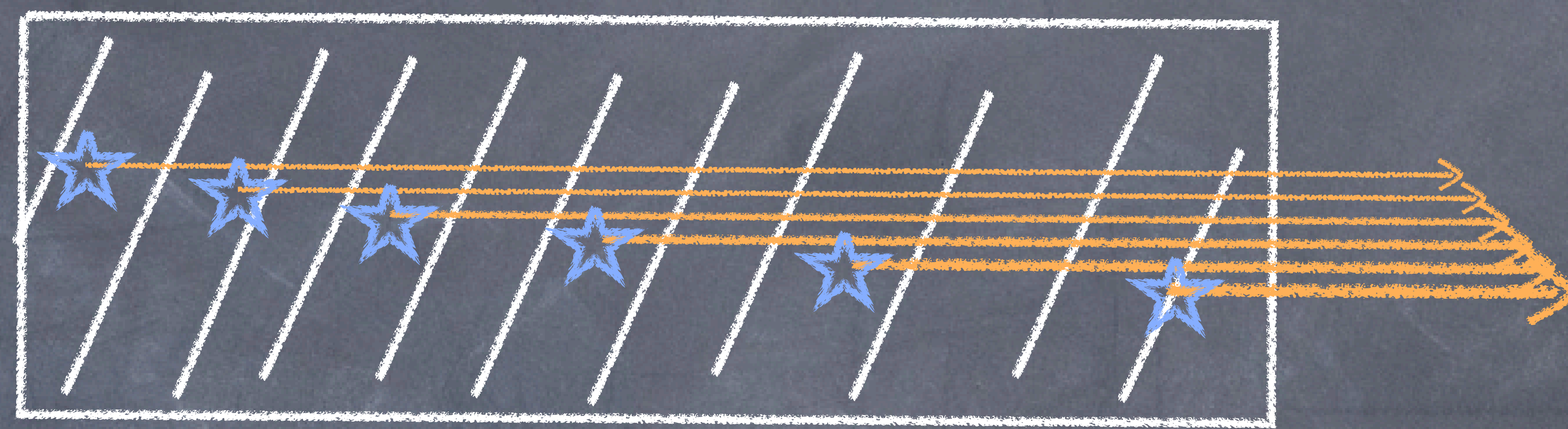
which looks better...

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau} + I_\nu = \underbrace{\frac{j_\nu}{4\pi\kappa_\nu} + \frac{\kappa_\nu^{\mathrm{sca}}}{\kappa_\nu}\Phi_\nu \equiv S_\nu}_{\text{the “source function”}}$$

without sources, we quickly see that

$$I(\tau) = I_0 e^{-\tau}$$

the canonical result that the **intensity decreases exponentially** with optical depth with sources and absorbers:



it's like each source is independently attenuated according to

$$I(\tau) = I_0 e^{-\tau}$$

(superposition of solution from different sources)

seems pretty simple
what's the big deal then?

that's in 1D, monochromatic...

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \hat{\mathbf{k}}\cdot\nabla I_\nu = \frac{1}{4\pi}\rho j_\nu - \rho\kappa_\nu^{\mathrm{abs}} I_\nu - \rho\kappa_\nu^{\mathrm{sca}} I_\nu + \rho\kappa_\nu^{\mathrm{sca}}\int \phi_\nu(\hat{\mathbf{k}},\hat{\mathbf{k}}') I_\nu(\hat{\mathbf{k}}')\mathrm{d}\Omega'$$

emissivity can depend on intensity at other wavelengths (like a heated blackbody...)

scattering couples different directions

This makes it hard!

do it numerically

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \hat{\mathbf{k}} \cdot \nabla I_\nu = \frac{1}{4\pi}\rho j_\nu - \rho\kappa_\nu^{\mathrm{abs}} I_\nu - \rho\kappa_\nu^{\mathrm{sca}} I_\nu + \rho\kappa_\nu^{\mathrm{sca}} \int \phi_\nu(\hat{\mathbf{k}}, \hat{\mathbf{k}}') I_\nu(\hat{\mathbf{k}}') \mathrm{d}\Omega'$$

- The intensity depends on 6 independent variables - position, direction, and wavelength! (and time too, in some cases)
- If you try to solve it with a normal finite-difference scheme (like a hydro code), you'll get nowhere!
- Must be smarter...

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \hat{\mathbf{k}}\cdot\nabla I_\nu = \frac{1}{4\pi}\rho j_\nu - \rho\kappa_\nu^{\mathrm{abs}} I_\nu - \rho\kappa_\nu^{\mathrm{sca}} I_\nu + \rho\kappa_\nu^{\mathrm{sca}}\int \phi_\nu(\hat{\mathbf{k}},\hat{\mathbf{k}}')I_\nu(\hat{\mathbf{k}}')\mathrm{d}\Omega'$$

- if the optical depth is large, the mean free path is small and photons scatter so many times they forget where they came from
  - the problem then reduces to a diffusion problem ("diffusion approximation")
- if the radiation is absorbed and re-emitted repeatedly, the radiation field at any point looks like the emission at that point
  - this is called "local thermal equilibrium" (LTE)
- now 3-D, one variable problem
- this is the case in, e.g., the interior of stars

$dT/dr \sim \kappa r^{-2} T^{-3}$

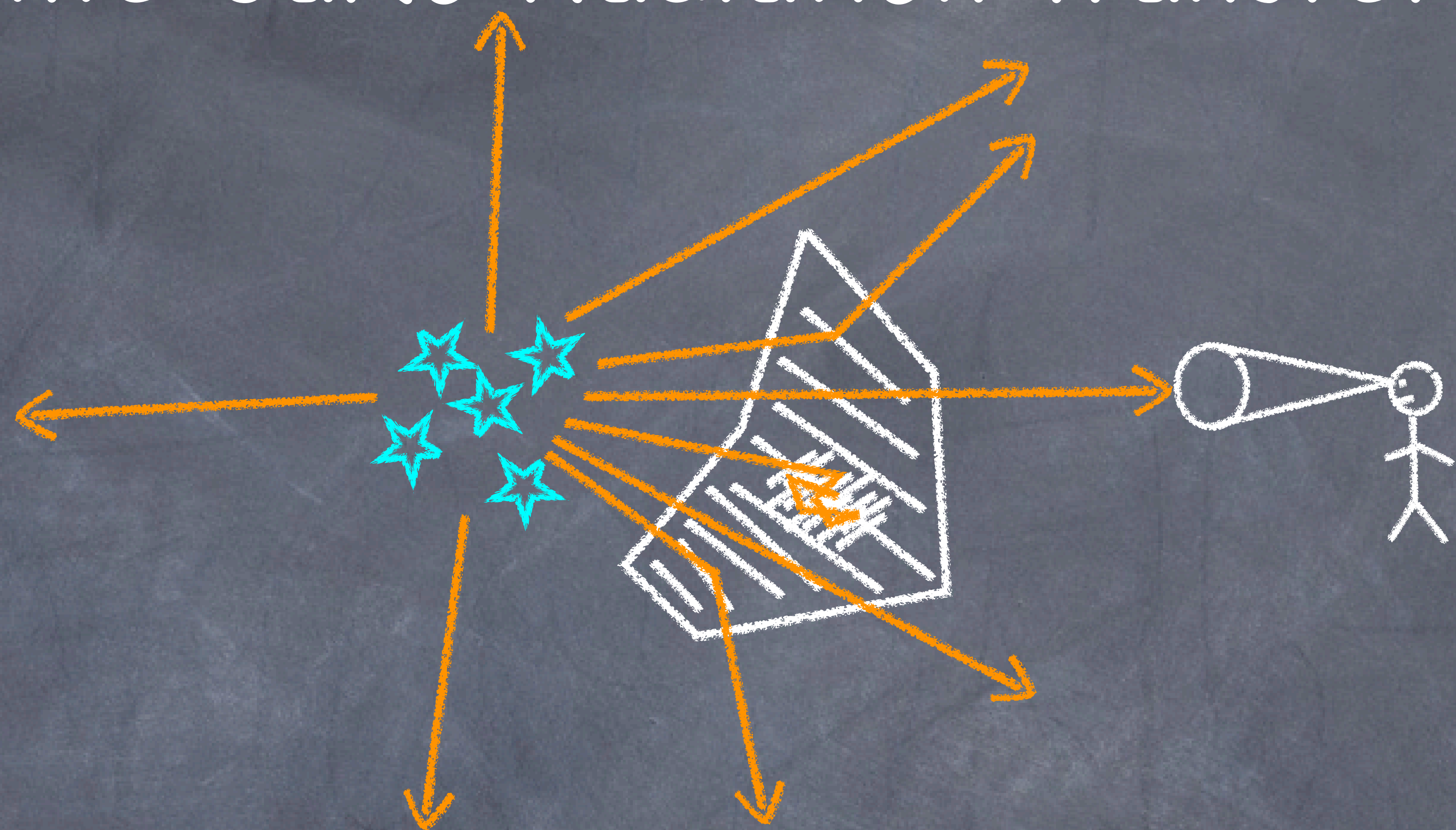
Other Approaches

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \hat{\mathbf{k}}\cdot\nabla I_\nu = \frac{1}{4\pi}\rho j_\nu - \rho\kappa_\nu^{\mathrm{abs}} I_\nu - \rho\kappa_\nu^{\mathrm{sca}} I_\nu + \rho\kappa_\nu^{\mathrm{sca}}\int \phi_\nu(\hat{\mathbf{k}},\hat{\mathbf{k}}')I_\nu(\hat{\mathbf{k}}')\mathrm{d}\Omega'$$

- Discretize directions and integrate along rays
  - one grid cell ("Short Characteristics")
  - to the edge ("Long Characteristics")
- Use moments of the intensity and write equations in terms of mean intensity and flux
  - "Variable Eddington Tensor" methods
- Sample the radiation field statistically
  - "Monte Carlo" methods

(note how suspiciously many papers are from Los Alamos...)

Monte Carlo Radiation Transfer



- solve RTE "like nature does"
- randomly emit "photons" from sources
- scatter/absorb them according to opacity
- make an image from rays that reach the observer
- as rays traverse the volume, they sample the radiation intensity distribution

Monte Carlo Radiation Transfer

- Advantages:
  - very general
  - easily handles arbitrary geometries or complicated media (scattering characteristics)
- Disadvantages:
  - solution contains Poisson noise
  - converges as $\sqrt{N}$, slowly
  - fails in the limit of very large τ
  - computationally expensive

MC example: photon propagation

Remember $I(\tau) = I_0 e^{-\tau}$?

this means that

$$\left|\frac{\mathrm{d}I}{I_0}(\tau)\right| = e^{-\tau}\mathrm{d}\tau$$

the probability of absorption is $P(\text{ absorption between } \tau, \tau + \mathrm{d}\tau) = e^{-\tau}$

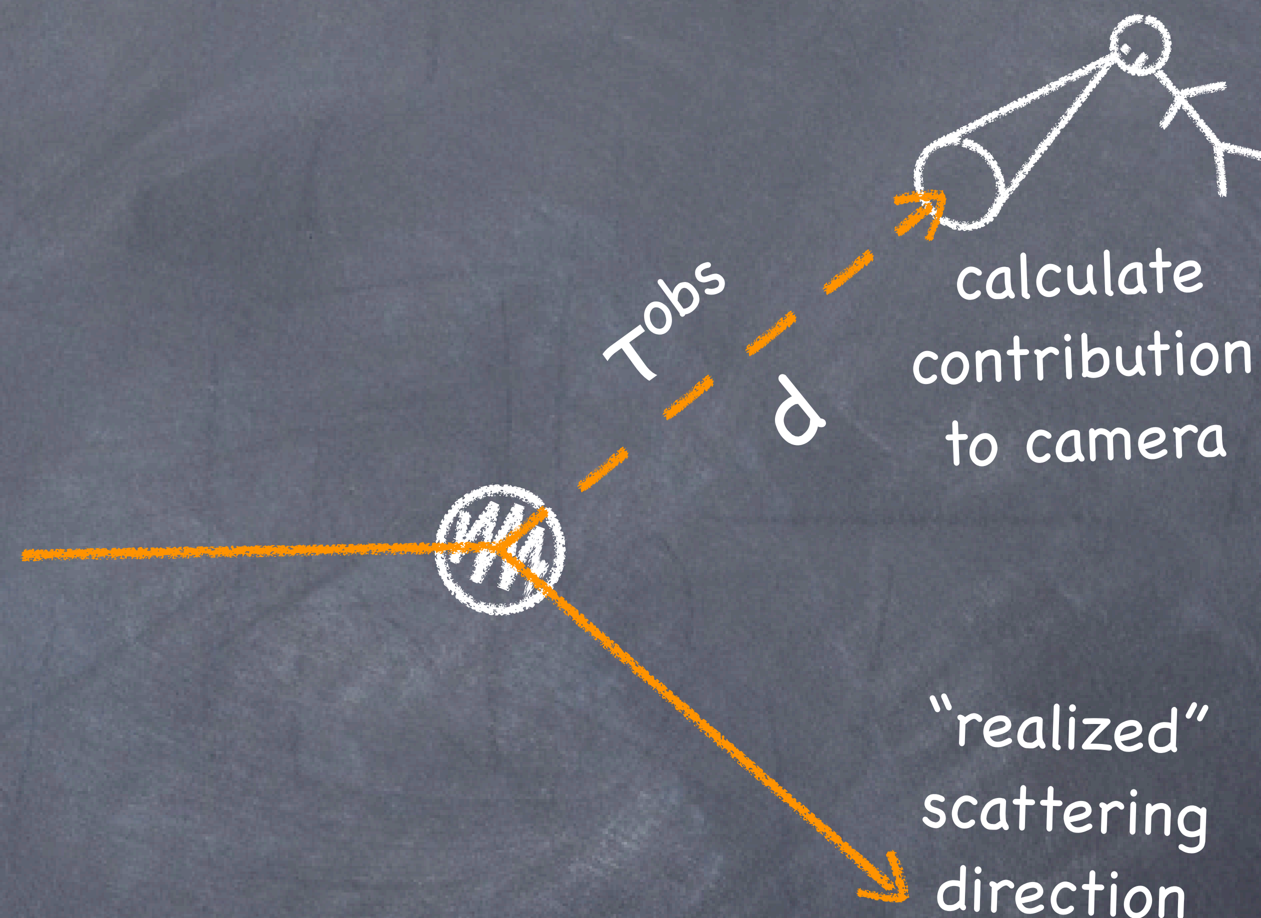
i.e.: the length a photon goes before it interacts is a random variable distributed as $e^{-\tau}$

Other processes

- in this way, sample the relevant processes
  - position and direction of emission
  - length of propagation
  - direction of scattering
- If you **can** apply an analytic solution instead of sampling it with MC, do it
  - example: use grain albedo to change the statistical "weight" instead of separately sampling absorption and scattering

“next event estimator” or “peel-off”

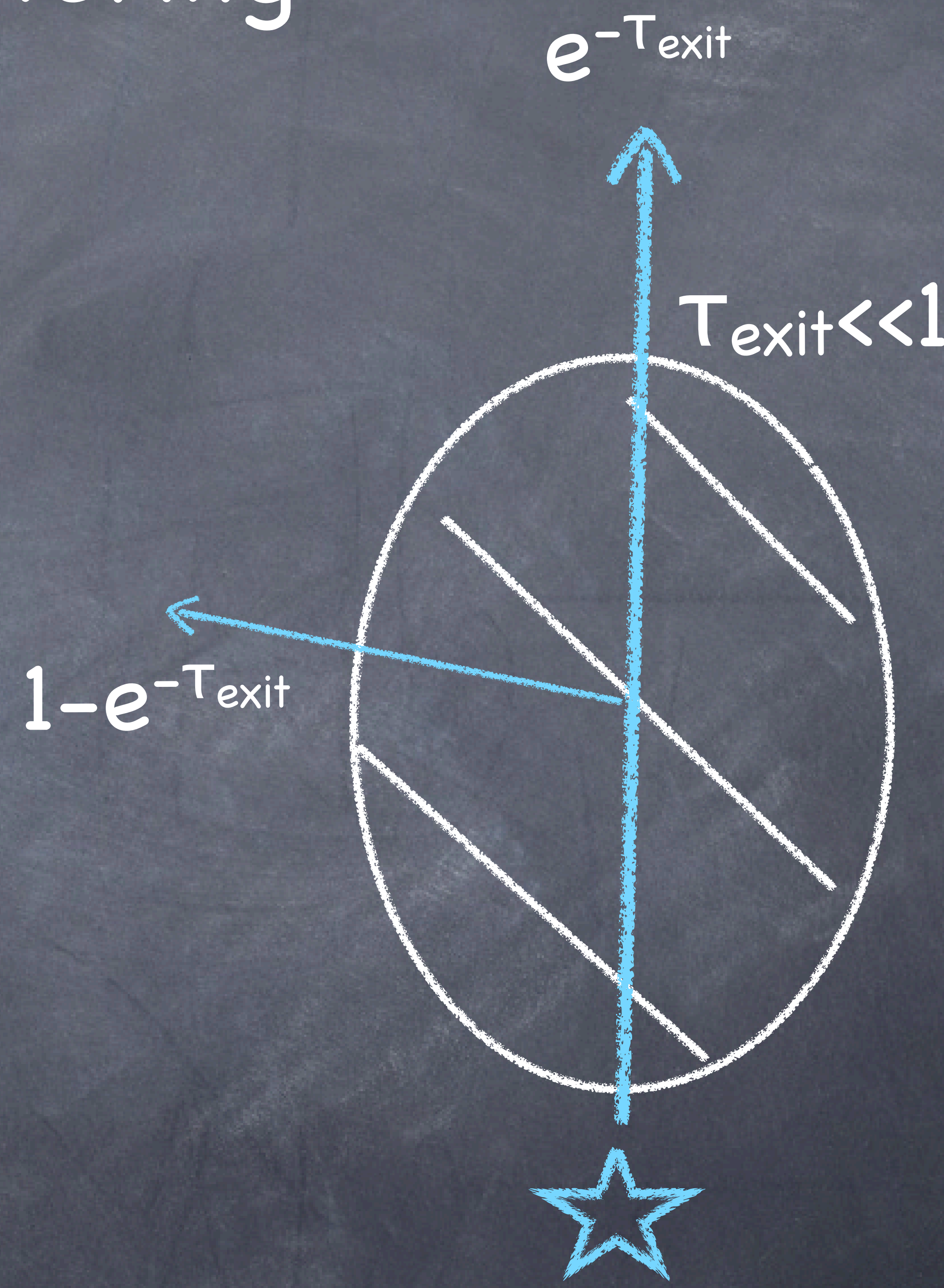
- Problem: if we randomly sample the ray random walk, practically none will reach the “camera”
- increase the efficiency (by a lot) by calculating the **probability** that a ray will reach the camera
- Example: scattering



$$F_{i,1} = L_i I_{i,1} e^{-\tau_{i,1}^{\rm obs}} \Phi_s(\hat{r}_{i,0}, \hat{r}_{i,1}^{\rm obs}) \frac{1}{d_{i,1}^2}$$

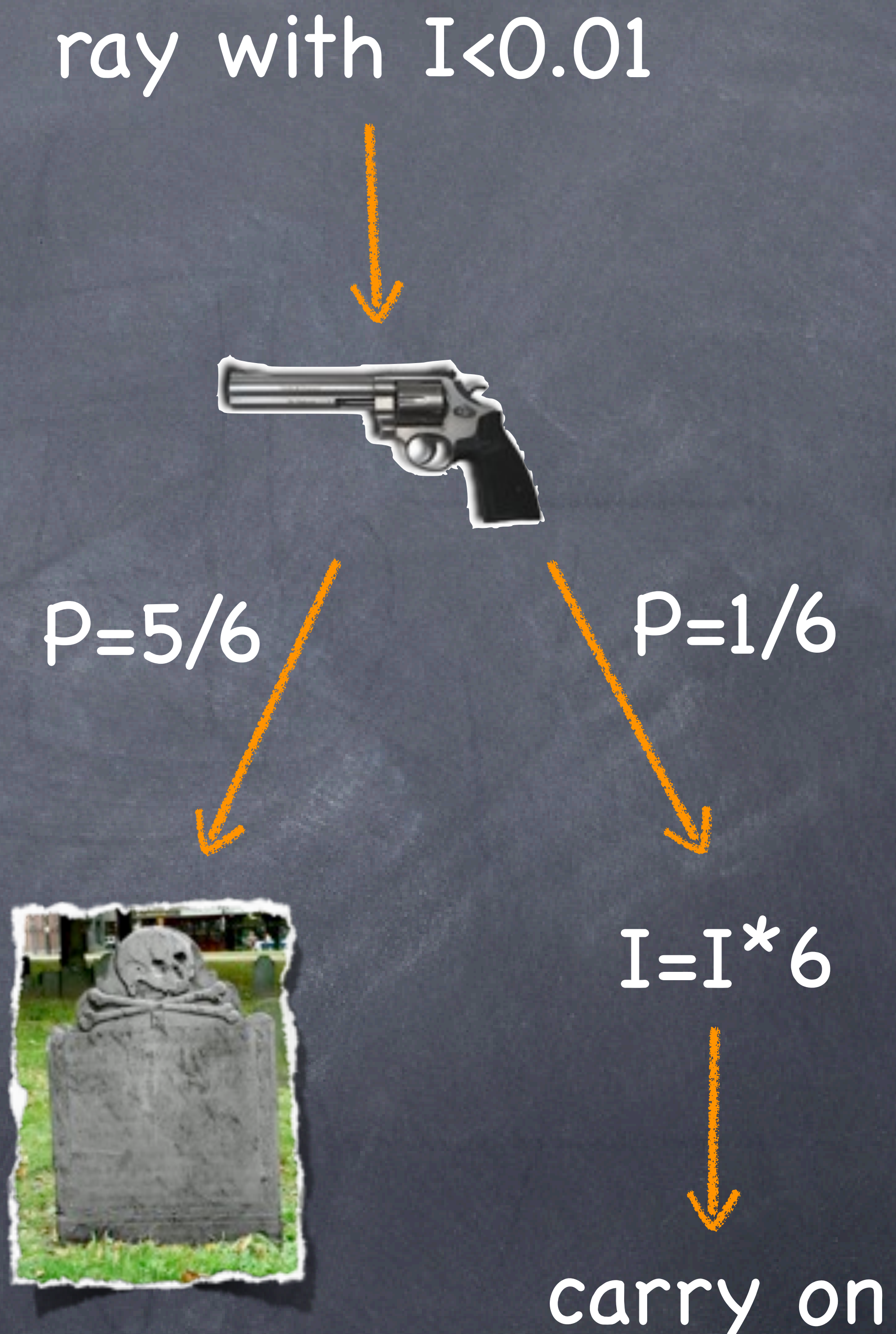
"Forced scattering"

- If medium is very optically thin, most rays pass through without scattering
  - Poor signal in the scattered light
- Can calculate analytically what fraction of the ray will leave and which will scatter **somewhere** on the way
- The location of the scattering event is then drawn from $[0,\tau_{exit}]$



“Russian Roulette”

- If a ray scatters many times, its intensity becomes low
- Don’t want to keep tracking a bunch of rays that won’t make much contribution
- But to preserve energy conservation, we can’t just drop the ray.



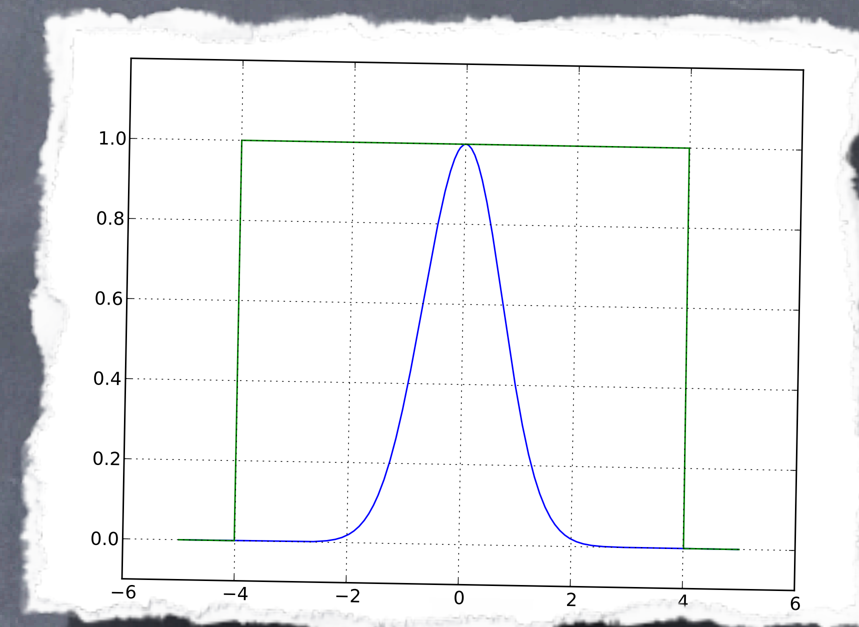
Polychromatic ray tracing

- All these distributions depend on wavelength, so a separate random walk is necessary for each wavelength
- In Sunrise, the computational cost of tracing the ray is dominated by walking the ray through the octree
- This means:
  - wavelength resolution is expensive!
  - Uncorrelated noise in spectra
- Can we do something more efficient?

Biased sampling

- **Biasing** – drawing from a different distribution than that sampled
- Suppose we want to sample f(x)
- We can do that while drawing from g(x)
- IF we also weight every sample x_i by $w_i = f(x_i)/g(x_i)$
- only requirement is that

  $g(x) > 0 \ \forall x$ where $f(x) > 0$



Can draw numbers from a gaussian distribution

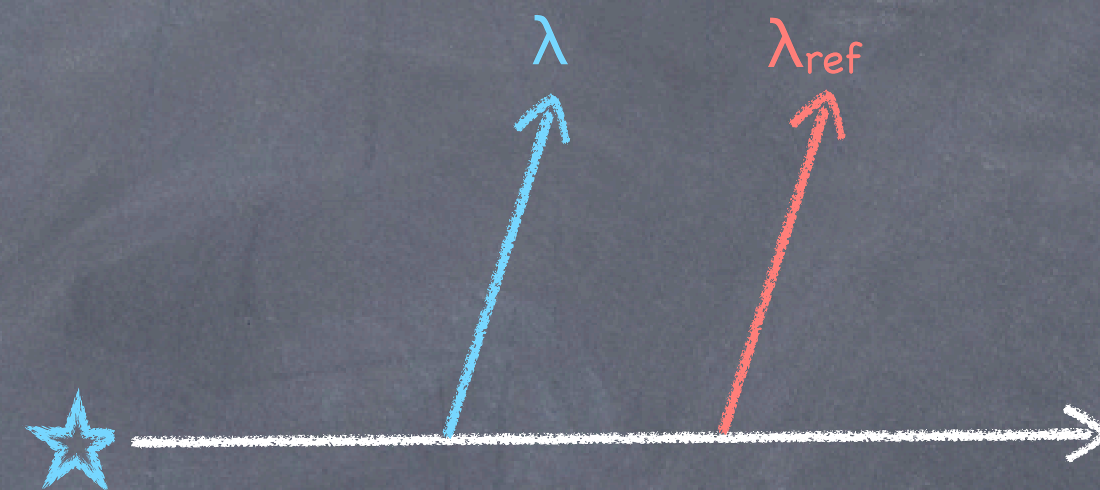
trivial

but:
what if we need to sample the wings??

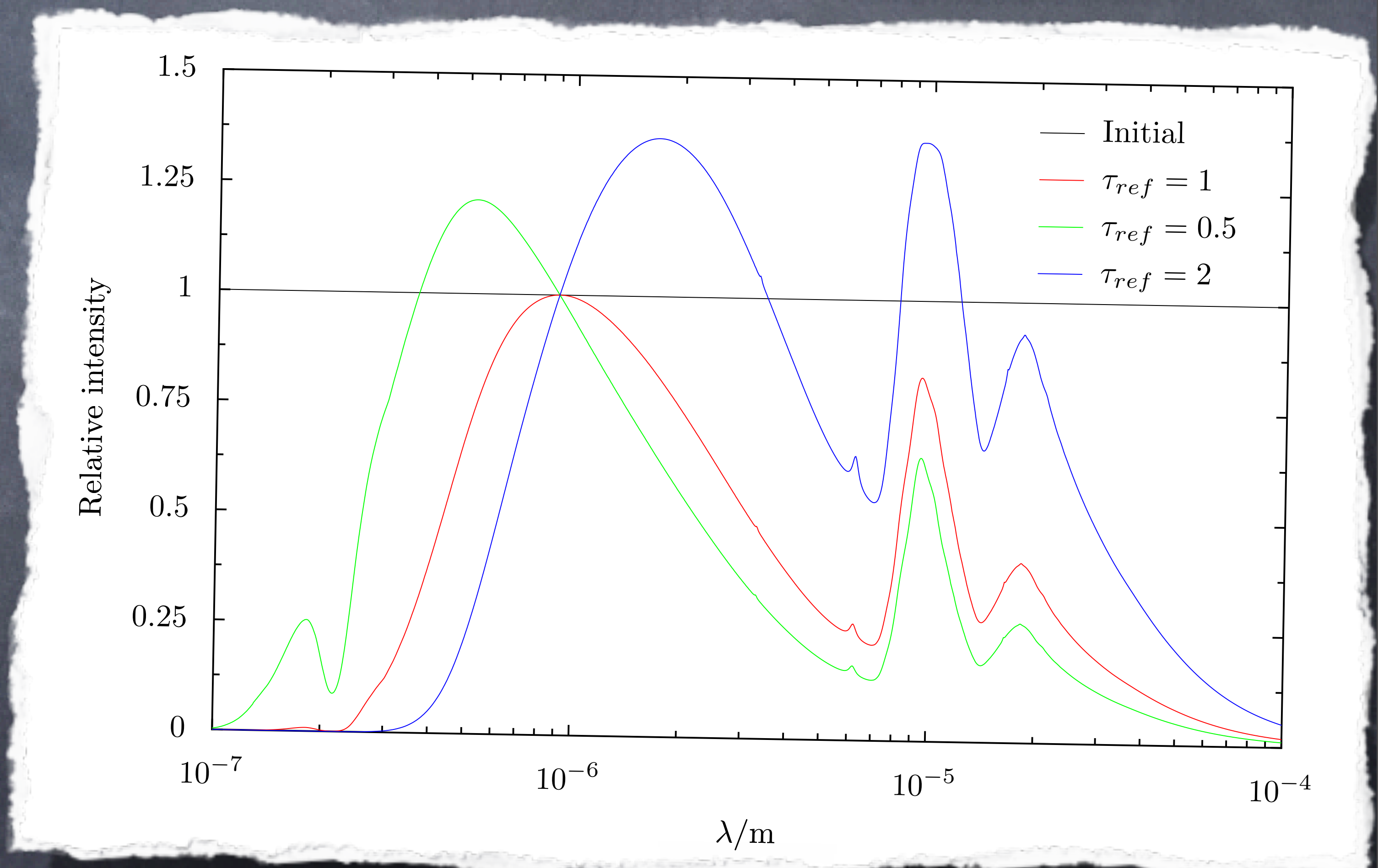
we can also draw numbers from a **uniform** distribution

by giving each sample a gaussian weight

Polychromatic ray tracing



- Mean free path varies with wavelength
- Can only draw scattering point correctly for one wavelength – λ_{ref}
- The other wavelengths are weighted according to the probability of them interacting at the drawn point
- Converges to correct distribution for all wavelengths



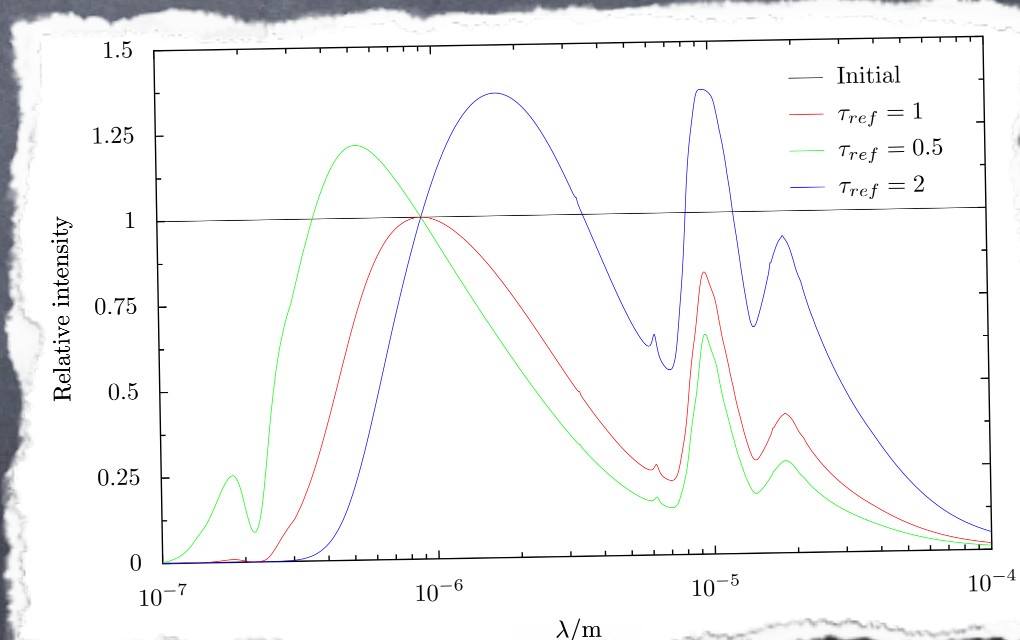
Polychromatic ray tracing

Probability of wavelength λ interacting at τ(λ) is

$$\mathrm{d}P\left[\tau(\lambda)\right] = \mathrm{e}^{-\tau(\lambda)}\mathrm{d}\tau(\lambda) = \mathrm{e}^{-(\tau(\lambda)/\tau_{\mathrm{ref}})\tau_{\mathrm{ref}}}\left[\frac{\tau(\lambda)}{\tau_{\mathrm{ref}}}\right]\mathrm{d}\tau_{\mathrm{ref}}$$

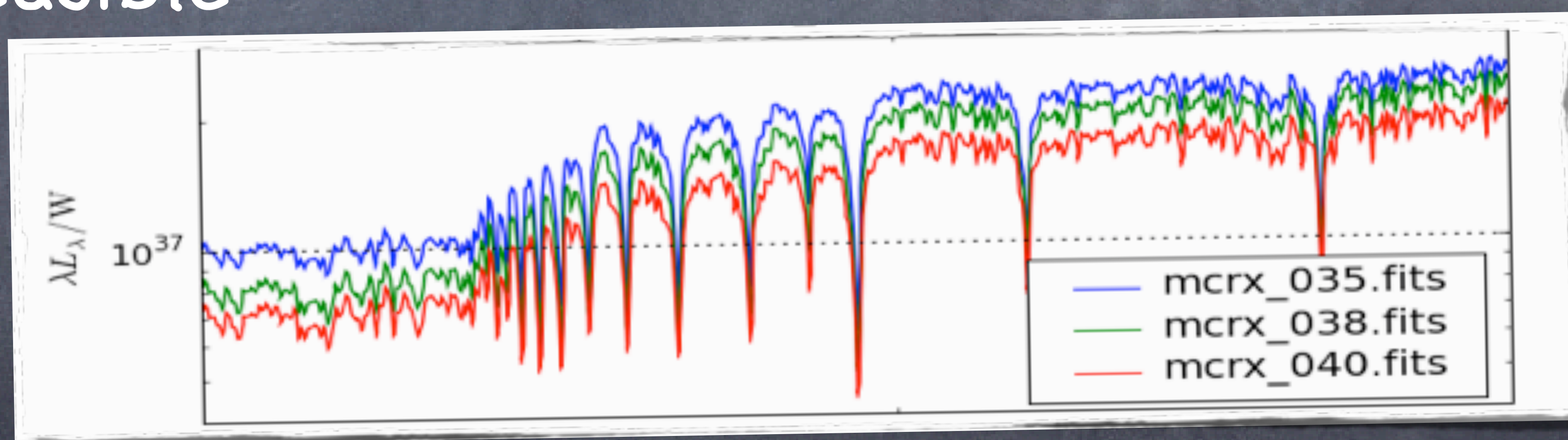
so if we sample it like $e^{-\tau_{ref}}\,d\tau$ we need to weight it by

$$w_\lambda = \frac{P\left[\tau(\lambda)\right]}{P\left[\tau_{\mathrm{ref}}\right]} = \mathrm{e}^{\tau_{\mathrm{ref}}-\tau(\lambda)}\left[\frac{\tau(\lambda)}{\tau_{\mathrm{ref}}}\right]$$



Polychromatic ray tracing

- Now each wavelength is not a separate random walk but instead just a vector operation - much faster!
- No (uncorrelated) noise between wavelengths
- Makes the very high wavelength resolution feasible



- Nothing's for free though...

Polychromatic ray tracing

Drawback:

$$w_\lambda = \frac{P\left[\tau(\lambda)\right]}{P\left[\tau_{\mathrm{ref}}\right]} = \mathrm{e}^{\tau_{\mathrm{ref}} - \tau(\lambda)} \left[\frac{\tau(\lambda)}{\tau_{\mathrm{ref}}}\right]$$

if $\tau(\lambda)$ very different from τ_{ref}

w can be large $\longrightarrow$ increased noise

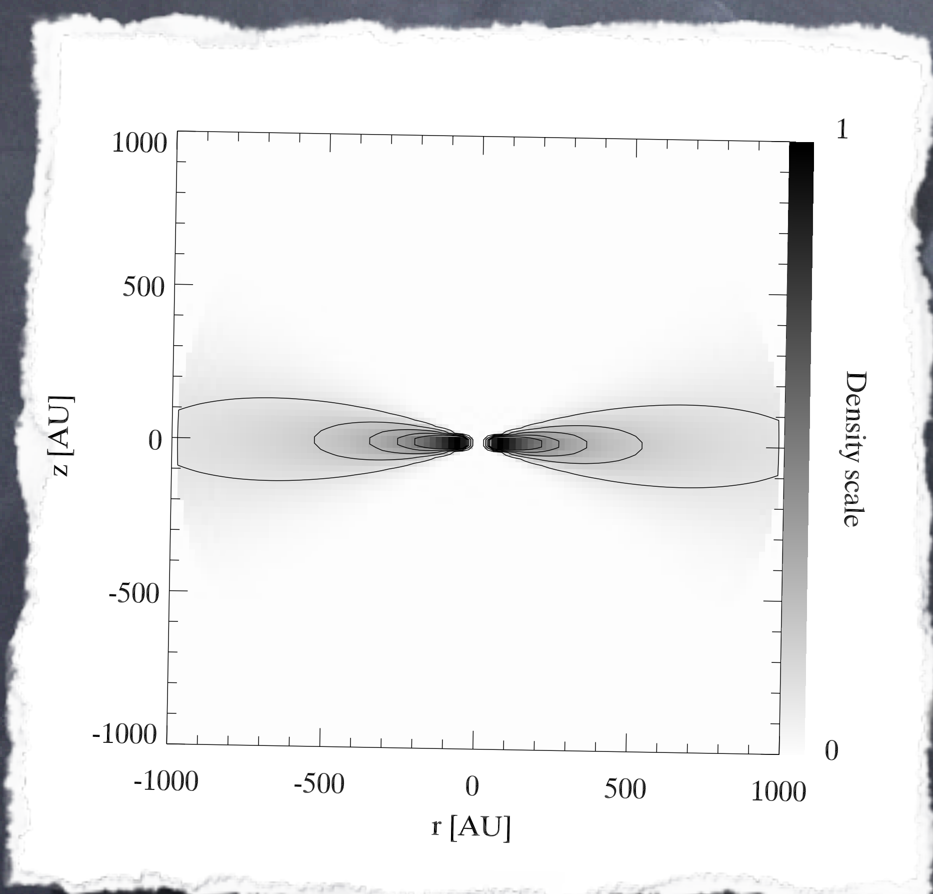
Bad situations:

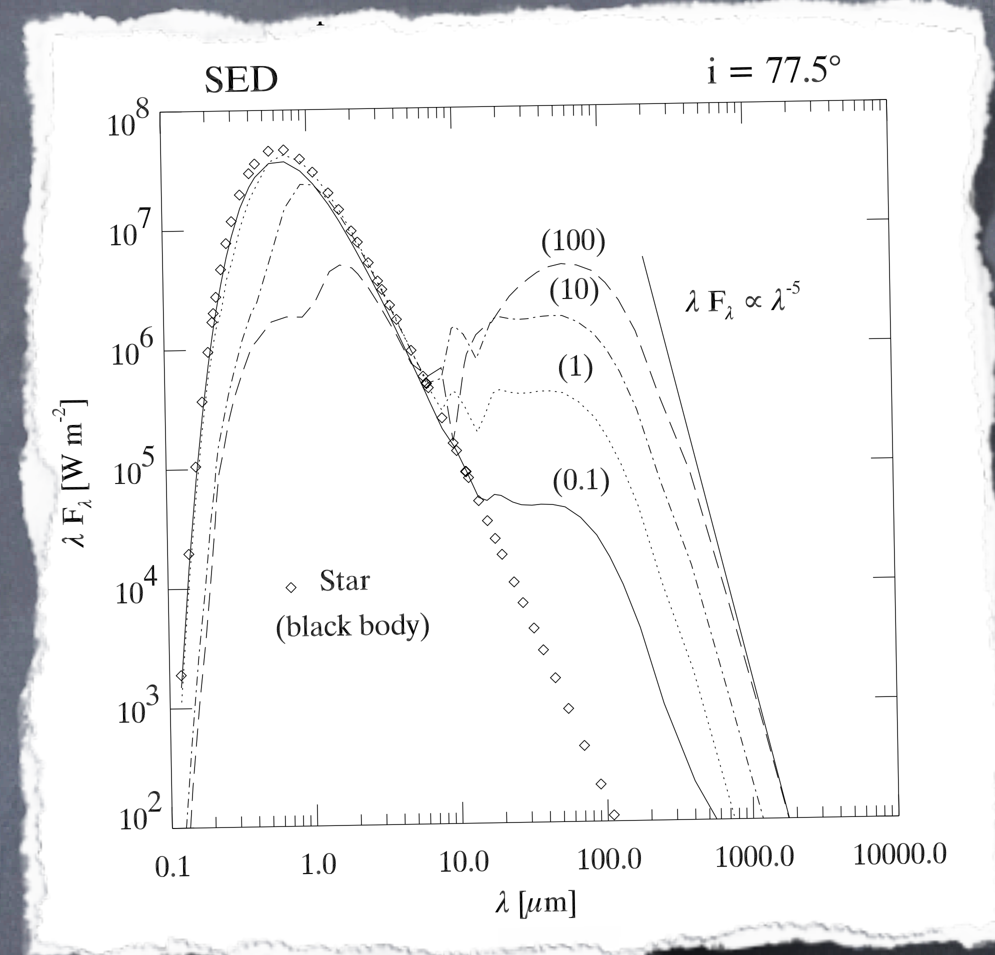
- very large optical depths
- rapidly changing opacity (e.g. lines!)

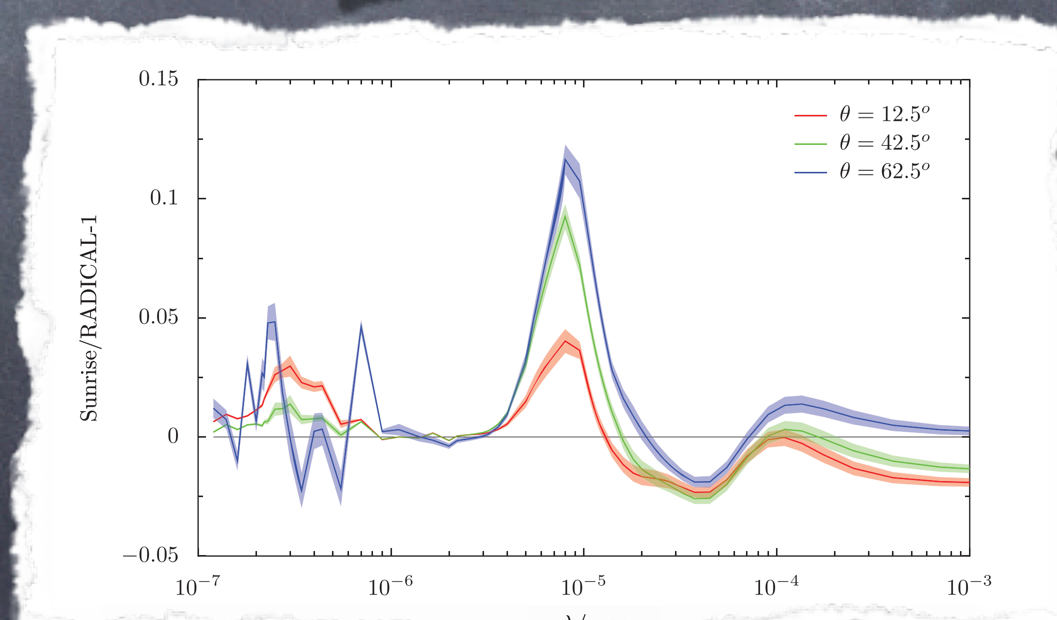
- Mitigated by splitting rays

Does this all work?

Pascucci et al. 2004
2D RT benchmark







The other codes did 50 calculations, polychromatic Sunrise did 1...

Intensity estimator

- Want to estimate the mean radiative intensity in the grid cells (for determining dust temperatures)
- Use the "path length estimator" (Lucy 99)
  - $J = \mathrm{sum}_i(I_i\ dl_i)/(4\pi V)$
  - I_i is the luminosity carried by ray i, dL_i the path length through the cell, V the cell volume

Parallelization

- This method is trivial to parallelize
- Each processor shoots its own ray, independent of every other
- Only need to worry about locking shared outputs: camera images and radiation intensities in cells
- With distributed memory, very different approach is needed: domain needs to be decomposed, rays need to be shifted from processor to processor as they travel

References

- Variable Tensor Methods:
  - Gnedin & Abel 2001, New Astronomy, 6, 437
- Long Characteristics/ray tracing
  - Abel & Wandelt 2002, MNRAS, 330, 53
- Short Characteristics:
  - Dullemond & Turolla 2000, A&A, 360, 1187
- Flux-limited diffusion
  - Levermore & Pomraning 1981, ApJ, 248, 321
- Monte Carlo
  - Jonsson 2006, MNRAS, 372, 2 (contains all references you want)
  - Dupree & Fraley 2002: "A Monte Carlo Primer"
  - Lux & Koblinger 1991: "Monte Carlo Transport Methods"
- Radiation Transfer Test cases
  - Pascucci et al. 2004, A&A, 417, 793
  - Iliev et al. 2006, MNRAS, 371, 1057