LA-UR- 11-02227

Approved for public release; distribution is unlimited.

Title:	Common Misconceptions in Monte Carlo Particle Transport
Author(s):	Thomas E. Booth
Intended for:	IRRMA-8, Kansas City, June 26 – July 1, 2011 (Industrial Radiation and Radioisotope Measurement Applications)

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the Los Alamos National Security, LLC for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

Form 836 (7/06)

meeting
IRRMA-8 KANSAS CITY 6/26 - 7/1/2011

Common Misconceptions in Monte Carlo Particle Transport

Thomas E. Booth

October 25, 2010 LA-UR 10-07691

1 Abstract

Monte Carlo particle transport is a well-established field, but a number of common misconceptions persist among many members of the Monte Carlo transport community. Some of these misconceptions directly impact current transport calculations and some are interesting to understand when designing new variance reduction techniques.

First, there is some confusion about the role of adjoints in forward Monte Carlo simulations. What are adjoints from a Monte Carlo perspective? It is common practice to bias a Monte Carlo calculation using adjoint information. Why? When does this practice work well and when does it not? What are the inherent dangers in basing weight windows on adjoints? The S_N codes typically only supply a free-flight adjoint. What is an entering collision adjoint and where would Monte Carlo codes use this adjoint? Second, S_N and Monte Carlo codes are being coupled together to obtain pulse height tallies. There are some inherent errors associated with this coupling that are often either not recognized, not understood, or simply ignored. Third, on a more theoretical level, many people intuitively interpret "statistical weight" as representing a number of particles. For many transport problems it is shown that this interpretation either severely restricts the type of Monte Carlo methods available or is simply wrong.

Common Misconceptions in Monte Carlo Particle Transport

Thomas E. Booth
Group XCP-4 Methods and Algorithms
Los Alamos National Laboratory
LA-UR-

IRRMA-8, Kansas City, June 26 – July 1, 2011

OUTLINE – Common Misconceptions

1. "The" transport equation describes all particle transport

2. Monte Carlo solves the transport equation

3. Monte Carlo and S_N pulse height tally estimates.

4. Monte Carlo and probability of initiation estimates.

5. Typical education bias: solving a problem is synonymous with solving an Eq.

OUTLINE – Common Misconceptions

6. "The" adjoint and importance functions.

7. Many ways to get a zero variance calculation.

8. Splitting concepts and information collection.

9. Zero variance, importance, and optimal weight windows.

10. List of other misconceptions.

11. Summary

"The" transport equation

The transport equation can be derived as a specific average over a Monte Carlo transport problem.

Monte Carlo "solves" the transport equation in the sense that the transport equation can (sometimes) predict the average Monte Carlo results.

Monte Carlo transport can also solve problems that the Boltzmann transport equation cannot.

Examples:
Pulse height estimates, coincidence estimates, probability of extinction estimates

The usual transport equation contains no information about the correlation between particles, because the equation derivation ignored correlation between particles. (A different transport equation could be derived using a different averaging process that did not average over the correlation.)

Monte Carlo codes (e.g. MCNP) can estimate tallies depending on correlation between particles (e.g. pulse height tallies) only because the Monte Carlo codes keep track of the correlation.

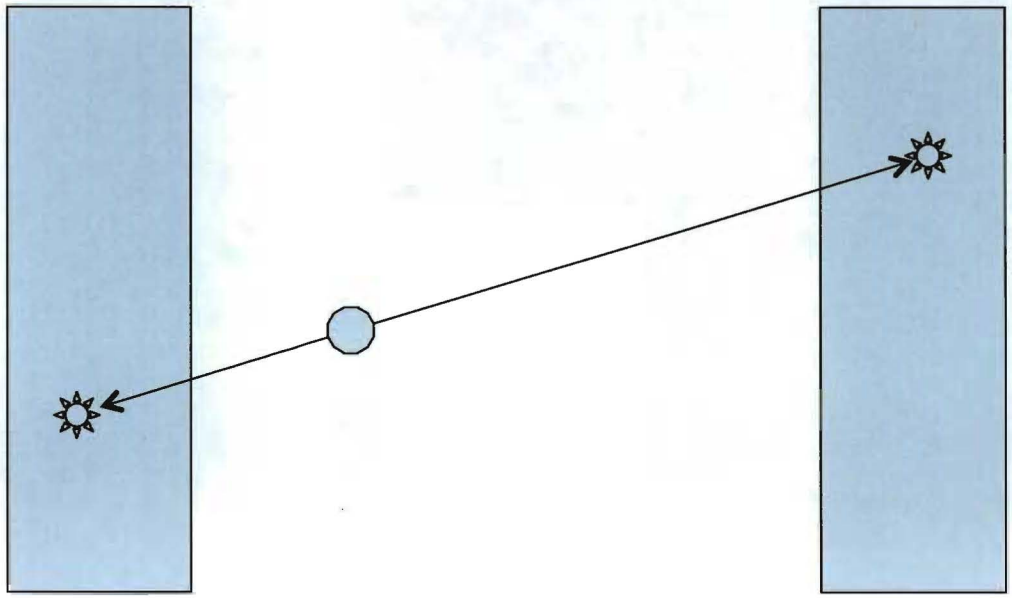
A Caution on Using S_N Transport Codes for Pulse Height Tallies

There are claims that S_N codes can be used to obtain pulse height tallies such as the MCNP F8 tally.

Sometimes these claims are even empirically validated by direct comparison with the MCNP F8 tally.

The problem is that the pulse height tally depends on correlation between particles that the S_N codes ignore.

Good comparisons with MCNP can be obtained only when it is unlikely that two correlated particles will both reach the detector. CAVEAT EMPTOR



Two 0.511 MeV photons from pair annihilation event absorbed in a pulse height detector must contribute 1 hit in the 1.022 MeV bin and not two hits in the 0.511 Mev bin.

Standard S_N codes do not calculate the required joint density of these two correlated particles.

Analog Monte Carlo transport codes can get pulse height tallies because the code knows that all particles from a given history are a correlated collection of particles.

Historically, variance reduction techniques were designed for tallies dependent only on individual particle densities and not joint densities.

Two choices:

1. Only do analog pulse height calculations (possible, but often prohibitively expensive), OR
2. Redefine the variance reduction to apply to collections of particles that tally.

Probability of Initiation Calculations (POI or POE=1-POI)

For a slightly supercritical system, the introduction of a stray neutron might produce a divergent neutron chain or it might not. What is the probability that a divergent chain occurs?

People have tried to use standard Boltzmann Monte Carlo transport concepts to solve this problem. Often, people erroneously come to the conclusion that it is impossible to get an exact Monte Carlo POI estimate.

POI calculations are even more problematical than pulse height tallies because they cannot be done even with an analog Monte Carlo approach.

Typical approach is to realize that a computer cannot score on a divergent chain because the chain never finishes.

Make some arbitrary definition (say N > 10000) that the chain has "diverged". Answers approach truth as N goes to infinity, but the computer time approaches infinity also.

An exact calculation is possible and efficient once one uses appropriate Monte Carlo concepts. Score on extinctions and use POI=1-POE. Use variance reduction to stochastically eliminate long (unimportant) chains.

Roulette long chains or importance sample chains to favor chains likely to terminate. Weight is assigned to chains.

An interesting thing to note is that an analog POE calculation is impossible. With variance reduction it is not only possible but can be very efficient if a good importance (not "Boltzmann transport") function is available. (Exponential convergence for one simple problem.)

Educational Training for Solving Problems

From junior high onward (10-15 years), solving a problem is almost synonymous with solving an equation.

The notion that Monte Carlo "solves" the Boltzmann transport equation seems similar to the notion that a basketball "solves" Newton's equation when making a basket.

Certainly the ball obeys Newton's equations, but it is a stretch to say that the ball "solves" the equation.

One could write a Monte Carlo transport code without ever having seen the transport equation.

A DETERMINISTIC PERSPECTIVE ON ADJOINTS

Most nuclear engineers encounter the adjoint equation shortly after encountering the transport equation.

1. An equation that is mathematically adjoint to the transport equation is derived.

2. The solution to the adjoint equation (the adjoint flux) at phase-space point P is then interpreted as the expected score from a unit weight particle at P.

A MONTE CARLO PERSPECTIVE ON ADJOINTS

For Monte Carlo purposes, it is sometimes useful to invert this process.

1. Define a particle's importance as the expected score produced by a unit weight particle.

2. Write an equation for the expected score and show that the equation is adjoint to the transport equation.

Step 2 is optional from a purely Monte Carlo viewpoint.

A MONTE CARLO PERSPECTIVE ON ADJOINTS

Note that the "expected score" is a very simple concept.

Monte Carlo can estimate the expected score without ever considering whether or not the expected score function is adjoint to some other function or not.

MCNP's weight window generator estimates the importance by simply keeping records of the total weight entering a region and the total score produced by that weight.

That is, MCNP estimates the average score or importance by simple averaging.

A MONTE CARLO PERSPECTIVE ON ADJOINTS

There are many kinds of importances

Two common importance functions are:

1. The "free-flight" importance is the expected score produced by a unit weight particle that is moving toward its next collision. (Spanier & Gelbard notation $\chi^*(P)$)

2. The "entering collision" importance is the expected score produced by a particle that is entering a collision. (Spanier & Gelbard notation $\psi^*(P)$)

A MONTE CARLO PERSPECTIVE ON ADJOINTS

One can take a particle at any stage of a Monte Carlo calculation and define an expected score.

For example, there is an expected score associated with a particle that has collided and is midway through the output sampling process.

Expected score for a particle that has collided with O^{16}, isotropically scattered $\Omega \rightarrow \Omega'$, and is awaiting the sampling of the output energy E'.

MANY WAYS TO GET A ZERO VARIANCE SOLUTION

Zero variance derivations almost always assume that the set of possible random walks is the same as the analog set of random walks and the only thing that changes is the probability of these random walks.

This is an extremely limited view of zero variance theory. Note that it is theoretically possible to get zero variance solutions in the presence of splitting, forced collisions, weight windows, etc. for which the set of possible random walks is NOT the analog set of walks.

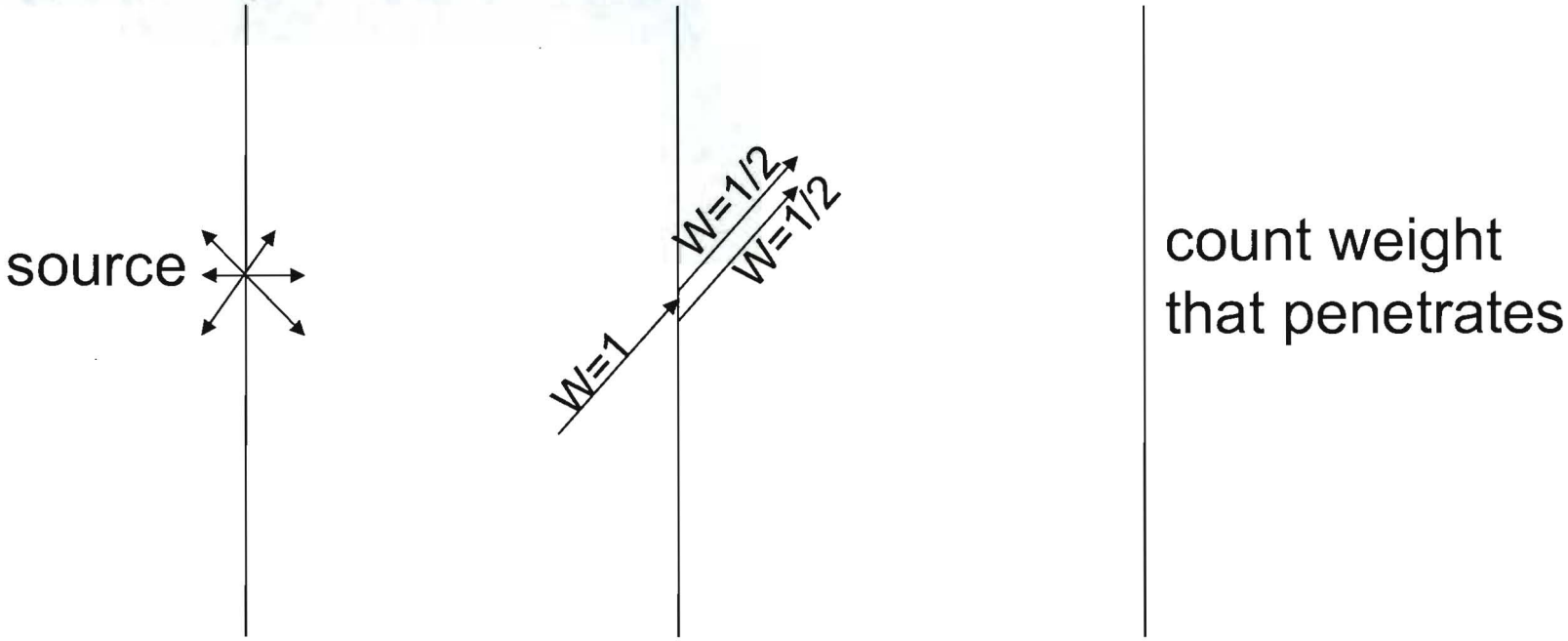
GENERAL WAY TO GET ZERO VARIANCE SOLUTIONS

If every random decision (sampling) is weighted by the expected score resulting subsequent to the sampling.

Note that the way most nuclear books and papers explain zero variance procedures is extremely limiting and equation focused.

Zero variance procedures are well-known in many fields besides transport, so it is useful to understand what the essence of a zero variance procedure is that connects all the different fields. *The common theme is not a particular equation …*

Splitting Concept

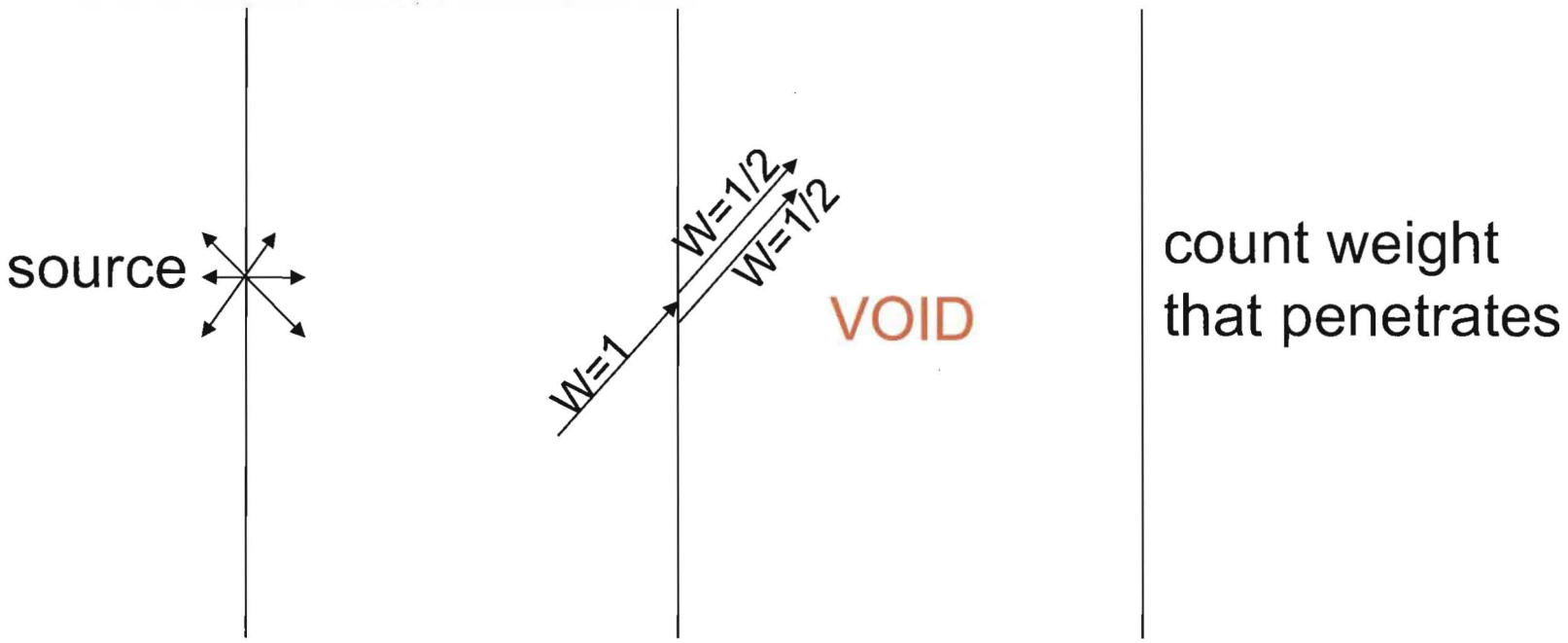


After split have two chances to penetrate.

Note variance reduction relies on the possibility that the split particles might do different things.

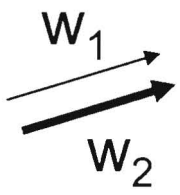
Possible scores 0, 1/2, 1

Splitting Concept (cont.)



If both particles always do the same thing then no new information is gained from the split. Might as well save the computer time and not follow two particles doing exactly the same thing.

Example of Inefficient Information Collection



particles at the same location
with extremely different weights

$$P_1=(x_1,y_1,z_1,v_{x1},v_{y1},v_{z1},t_1)= (x_2,y_2,z_2,v_{x2},v_{y2},v_{z2},t_2)=P_2$$

Suppose: $w_1=0.0001$ and $w_2=1$

Computer time to simulate events (e.g. collisions) is independent of weight.

Particle 2, on average, contributes 10,000 times as much to the score as particle 1.

Suppose: $w_1=0.0001$ and $w_2=1$

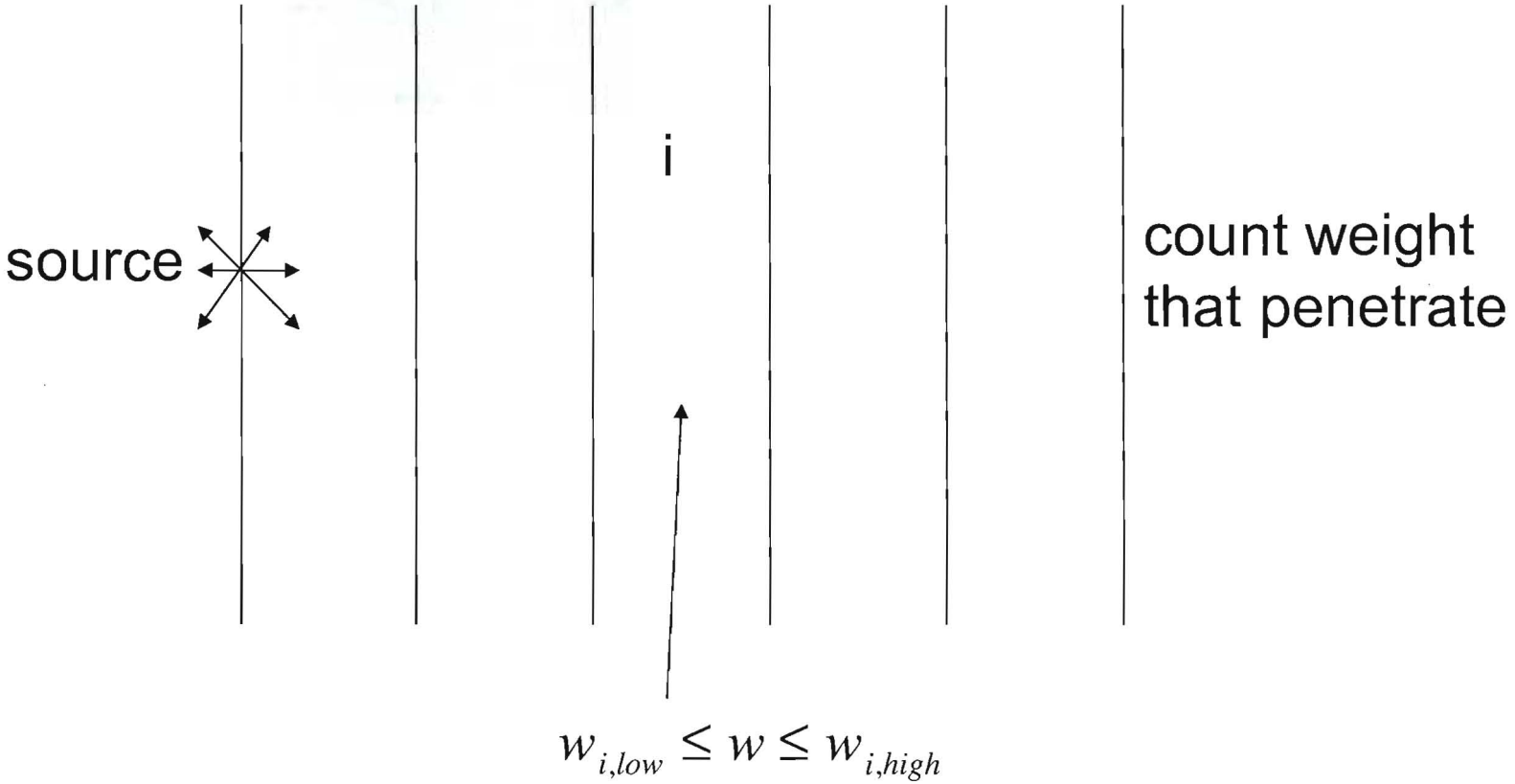
Computer time to simulate events (e.g. collisions) is independent of weight.

Particle 2, on average, contributes 10,000 times as much to the score as particle 1.

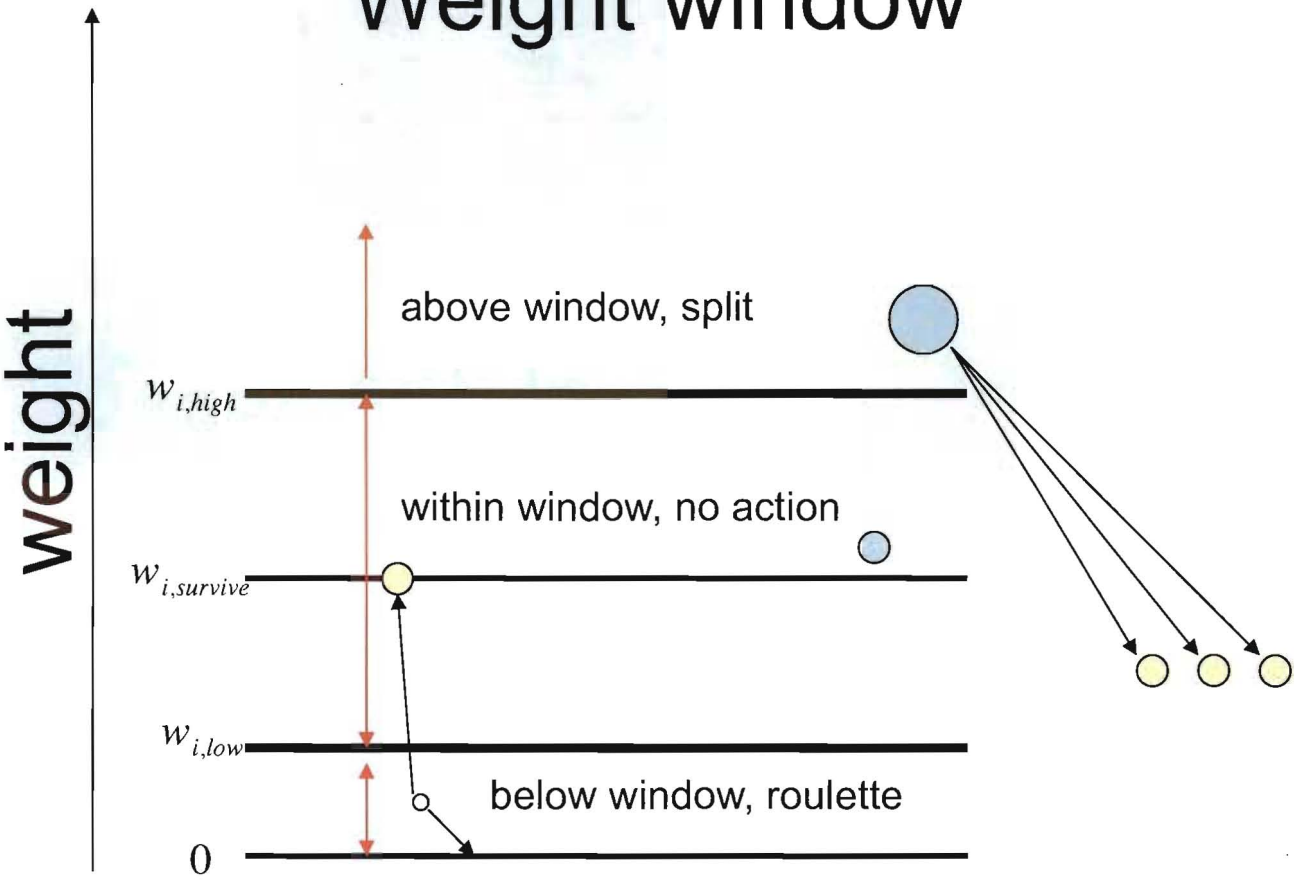
Makes no computational sense to spend much computer time on particle 1.

Particles in the same location should have roughly the same weights in an efficient calculation.

Weight Windows



Weight window



$$w_{i,high} = 5w_{i,low}$$

usually something like:

$$w_{i,survive} = 3w_{i,low}$$

The weight window forces each particle to have roughly the same weight, say,

$$w_{i,low} \le w \le 5w_{i,low}$$

but how should $w_{i,low}$ be chosen?

Typical rationale for choosing $w_{i,low}$:
Each particle should have roughly the same expected score, say C, independent of location P.

Let $M_1(P)$= expected score for a unit weight particle at P.
The expected score for a particle of weight w is w $M_1(P)$.
Thus want:

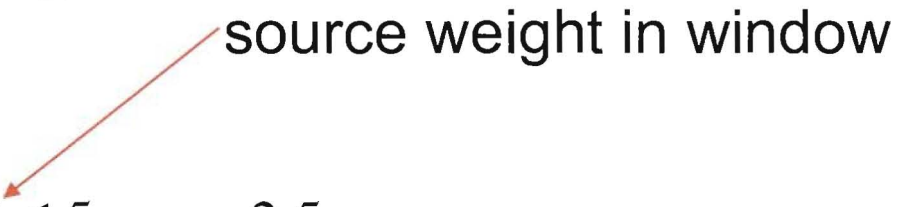
$$w_{low}(P)M_1(P) = C \qquad \text{or} \qquad w_{low}(P) = \frac{C}{M_1(P)}$$

All that is left is to specify C.

For this simple problem, unit weight particles are started at source point P_s and one chooses C so that source particles start within the window. For example, if

$$C = \frac{1}{2} M_1(P_s)$$

$$w_{low}(P_s) = \frac{C}{M_1(P_s)} = \frac{1}{2}$$



$$0.5 = w_{low} \leq 1 \leq 5 w_{low} = 2.5$$

$M_1(P)$ is sometimes obtained stochastically (e.g. the weight window generator in MCNP) and sometimes obtained by deterministically solving an equation for $M_1(P)$ (e.g. the adjoint S_N equations).

This procedure for obtaining a weight window works well in many (but not all) instances.

Note from statistical theory that the variance involves the first and second moments of the score distribution:

$$\sigma^2(P,w) = M_2(P,w) - M_1^2(P,w) = M_2(P,w) - w^2 M_1^2(P)$$

Note that M_1 is the same, independent of what variance reduction techniques (e.g. splitting) are used.

Minimizing the variance depends on minimizing M_2.

The weight windows are normally set using the adjoint, or expected score, function M_1 and ignoring the function M_2 that one would like to minimize.

This seems pretty odd? Why should focusing on the average particle behavior, M_1, often minimize M_2?

Why does this work at all?

Apparent Reasoning

1. Zero variance solutions are possible, in principle, using special zero variance procedures.

2. Zero variance procedures use adjoint (M_1) information

3. Then a miracle occurs.

4. Therefore, optimum low variance solutions using non-zero variance procedures (e.g. weight windows) can be obtained using adjoint information.

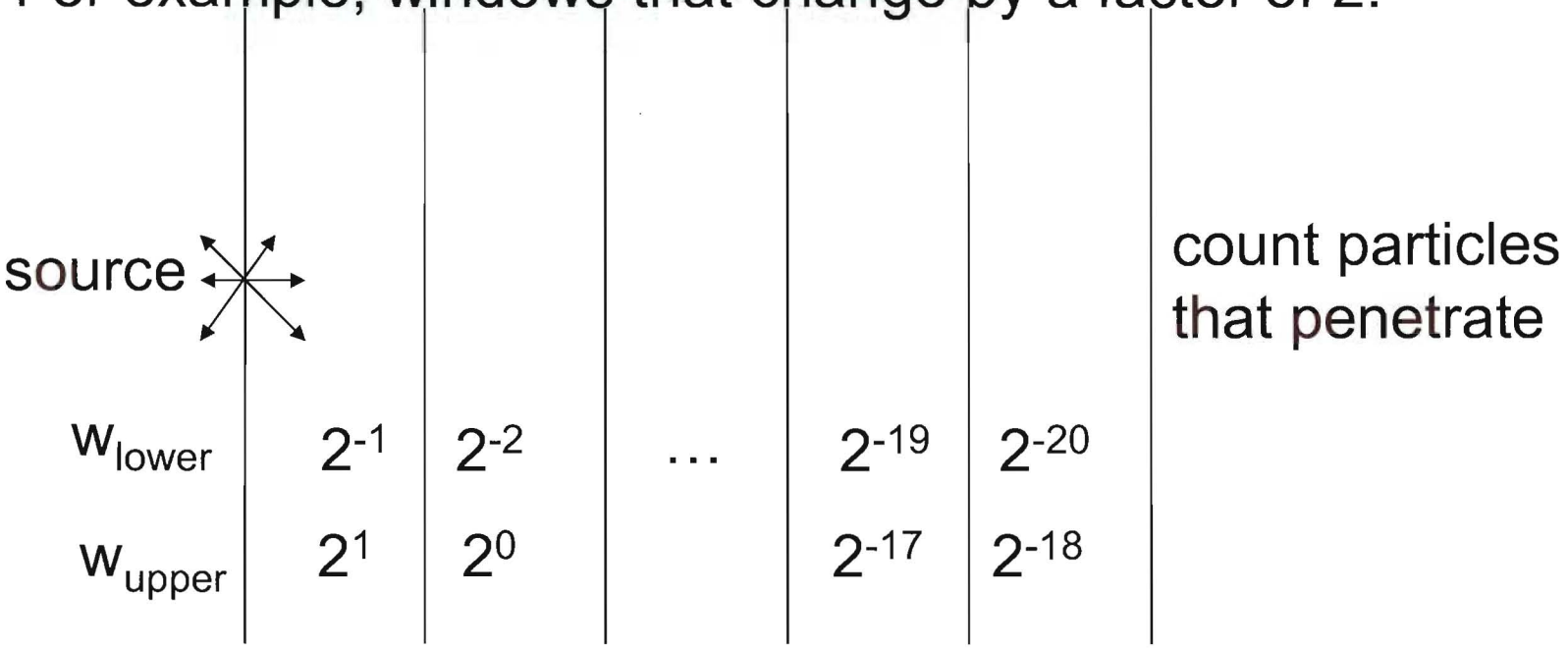
Consider a hypothetical zero variance solution.

Every source particle contributes the mean score M_1. The particles that contribute most to M_1 are the same particles that contribute most to M_2.

If a Monte Carlo simulation can be arranged so that most of the mean and most of the second moment result from the same particles, then setting the variance reduction using only M_1 typically works well.

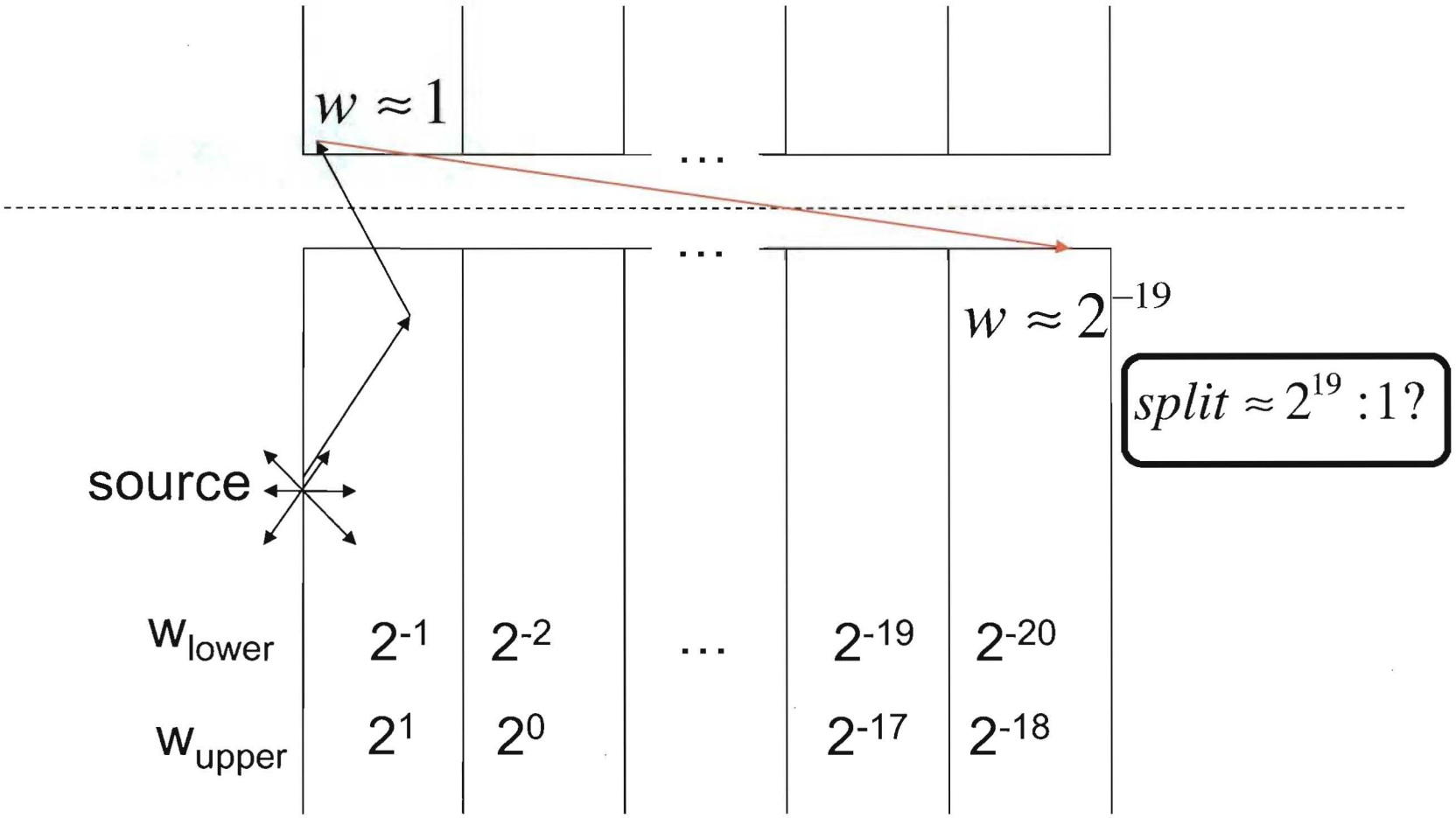
Setting weight windows using the adjoint works well when no large splits are required.

For example, windows that change by a factor of 2.



Problems with adjoint window setting

Cylindrical symmetry about void duct axis



Four points about a $split \approx 2^{19} : 1$

1. Splitting collects information when the split particles do different things.

2. There aren't really 2^{19} different types of things to do.

3. Most of the variance in this problem is due to the poor sampling of particles going up the duct.

4. The variance associated with the poor sampling up the duct has already occurred. This variance cannot be removed even with an infinite split.

Setting weight windows using the adjoint does not work well when large splits are required.

Note that for this problem the very rare particle that goes directly up the duct may have comparatively little impact on the mean, but a huge impact on the variance. Could be 99% of the variance but 1% of the mean.

Setting variance reduction according to what happens to the *typical* particles that contribute to the mean is not appropriate when the variance is driven by very *atypical* particles.

List of Other Common Misconceptions: Theoretical Interest

1. Weight represents a number of particles. Reality: weight is a tally multiplier. *Sometimes* represents a number of particles, sometimes not.
2. Weight > 0. (Actually, can be negative or even complex-valued.)
3. A necessary and sufficient condition for unbiased estimates is that the particle density is preserved. (This is neither necessary nor sufficient.)
4. A particle has one weight. (Actually, the particle can have many weights for many different reasons.)
5. Monte Carlo particle transport need *not* be Markovian. It is often presented that way because analog Monte Carlo is Markovian.

Summary

1. Focus on solving problems, not equations.
2. Use equations carefully.
   a. use the *correct* equation
   b. understand the *limitations* of the equation
3. Most Monte Carlo theory is for weight-independent transport; the math is easier. Most Monte Carlo codes do weight-dependent transport.

A Transport Process Approach to Understanding Monte Carlo Transport Methods:
***mcnp**-green.lanl.gov/publication/pdf/LA-UR-04-1426.pdf*