Monte Carlo Modelling of Many-Body Decays

Stephen Brooks

January 9, 2006

1 Introduction

The original implementation of particle decays in the *Muon1* tracking code only admitted two-body processes. These are easily dealt with by assuming the rest frame of the parent particle and producing two decay products with equal and opposite isotropically-selected momenta and magnitude such as to conserve energy. The problem of an n-body decay subtracts four constraints from the $3n$ degrees of freedom of the decay products: three for overall conservation of momentum and one for energy. The final states can therefore be found on a $(3n-4)$-dimensional manifold in the multi-particle phase space. This is two-dimensional for two-body decays (and can be made into a sphere by the method and symmetry argument above), but even for three-body decays it is already five-dimensional and rotational symmetry can remove at most three of these.

Faced with this problem of a complicated surface to generate the distribution on, a Monte Carlo routine that covers the right surface with points (at some known density) is obtained first and then the density is corrected afterwards by rejecting a calculated proportion of them and re-running the original routine as necessary (this is known as *rejection sampling*). Formally, the first routine generates the primal probability density $F(\mathbf{p}_i)$ on the feasible manifold, then any other desired distribution $f(\mathbf{p}_i)$ can be derived from this by only accepting each primal point with probability $\propto f(\mathbf{p}_i)/F(\mathbf{p}_i)$. Note that this should not be larger than 1, so the ratio must be bounded above and scaled accordingly.

2 The Probability Density

From the 2004 Particle Physics Booklet, section 38.4, the probability of a particle with 4-momentum P decaying into n particles of 4-momenta $p_1, p_2 \ldots p_n$ is governed by

$$\begin{aligned}\mathcal{P} &\propto \delta^4(P - \sum_{i=1}^{n} p_i) \prod_{i=1}^{n} \frac{\mathrm{d}^3\mathbf{p}_i}{E_i} \\ &= \delta(E - \sum_{i=1}^{n} E_i)\delta^3(\mathbf{P} - \sum_{i=1}^{n} \mathbf{p}_i) \prod_{i=1}^{n} \frac{\mathrm{d}^3\mathbf{p}_i}{E_i}.\end{aligned}$$

Here and elsewhere units are such that $c = 1$.

The delta functions can be multiplied into the main coefficient using the following rule, where f is a function with one root $f(a) = 0$:

$$\delta(f(x))g(x) = \frac{g(a)}{|f'(a)|}\delta(x - a).$$

Applying this three times, considering the probability expression as a function of p_n^x, p_n^y and p_n^z in turn, reduces it to:

$$\frac{\mathrm{d}^{3n}\mathcal{P}}{\mathrm{d}^3\mathbf{p}_1\mathrm{d}^3\mathbf{p}_2\ldots\mathrm{d}^3\mathbf{p}_n} \propto \left[\delta(E-\sum_{i=1}^{n}E_i)\prod_{i=1}^{n}\frac{1}{E_i}\right]\delta^3\left(\mathbf{p}_n-(\mathbf{P}-\sum_{i=1}^{n-1}\mathbf{p}_i)\right),$$

which when integrated over $\mathbf{p}_n$ gives

$$\frac{\mathrm{d}^{3n-3}\mathcal{P}}{\mathrm{d}^3\mathbf{p}_1\mathrm{d}^3\mathbf{p}_2\ldots\mathrm{d}^3\mathbf{p}_{n-1}} \propto \delta(E-\sum_{i=1}^{n}E_i)\prod_{i=1}^{n}\frac{1}{E_i},$$

after which the substitution $\mathbf{p}_n = \mathbf{P}-\sum_{i=1}^{n-1}\mathbf{p}_i$ should be used to determine the last momentum and E_n.

This needs to be integrated one more time to remove the last delta function, so a degree of freedom will be taken from $\mathbf{p}_{n-1}$. Let $\mathbf{p}_* = \mathbf{P}-\sum_{i=1}^{n-2}\mathbf{p}_i$ so that $\mathbf{p}_{n-1}+\mathbf{p}_n=\mathbf{p}_*$ to conserve 3-momentum. Now perform a Lorentz transformation into a primed frame where $\mathbf{p}'_*=\mathbf{0}$, so that $\mathbf{p}'_{n-1}=-\mathbf{p}'_n=p\hat{\mathbf{p}}$. Integrating over $p\geq 0$ will remove the delta function, leaving the two degrees of freedom in $\hat{\mathbf{p}}$ parameterising the remaining freedom of these last two momenta.

Starting with the expression written in the primed frame:

$$\frac{\mathrm{d}^{3n-3}\mathcal{P}}{\mathrm{d}^3\mathbf{p}'_1\mathrm{d}^3\mathbf{p}'_2\ldots\mathrm{d}^3\mathbf{p}'_{n-1}} \propto \delta(E'-\sum_{i=1}^{n}E'_i)\prod_{i=1}^{n}\frac{1}{E'_i}$$

...change $\mathbf{p}'_1$ through $\mathbf{p}'_{n-2}$ back to the original frame using $\mathrm{d}^3\mathbf{p}'_i=(E'_i/E_i)\mathrm{d}^3\mathbf{p}_i$:

$$\frac{\mathrm{d}^{3n-3}\mathcal{P}}{\mathrm{d}^3\mathbf{p}_1\mathrm{d}^3\mathbf{p}_2\ldots\mathrm{d}^3\mathbf{p}_{n-2}\mathrm{d}^3\mathbf{p}'_{n-1}} \propto \delta(E'-\sum_{i=1}^{n}E'_i)\frac{1}{E'_{n-1}E'_n}\prod_{i=1}^{n-2}\frac{1}{E_i}$$

...and then use the substitution, which implies $\mathrm{d}^3\mathbf{p}'_{n-1}=p^2\mathrm{d}p\mathrm{d}^2\hat{\mathbf{p}}$, to get:

$$\frac{\mathrm{d}^{3n-4}\mathcal{P}}{\mathrm{d}^3\mathbf{p}_1\mathrm{d}^3\mathbf{p}_2\ldots\mathrm{d}^3\mathbf{p}_{n-2}\mathrm{d}^2\hat{\mathbf{p}}} \propto \delta(E'-\sum_{i=1}^{n}E'_i)\frac{p^2\mathrm{d}p}{E'_{n-1}E'_n}\prod_{i=1}^{n-2}\frac{1}{E_i}.$$

Now that varying p is the concern, a "constant" part may be removed from the argument of the delta function

$$\delta(E'-\sum_{i=1}^{n}E'_i)=\delta(E'_*-(E'_{n-1}+E'_n)),$$

where $E'_*=E'-\sum_{i=1}^{n-2}E'_i$. The modulus of the derivative of this argument:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}p}(E'_{n-1}+E'_n) &= \frac{\mathrm{d}}{\mathrm{d}p}\left(\sqrt{p^2+m_{n-1}^2}+\sqrt{p^2+m_n^2}\right)\\
&= \frac{p}{\sqrt{p^2+m_{n-1}^2}}+\frac{p}{\sqrt{p^2+m_n^2}}\\
&= p\left(\frac{1}{E'_{n-1}}+\frac{1}{E'_n}\right)=p\frac{E'_{n-1}+E'_n}{E'_{n-1}E'_n},
\end{aligned}$$

...must be divided out when the argument of the delta function is linearised:

$$\begin{aligned}
\frac{\mathrm{d}^{3n-4}\mathcal{P}}{\mathrm{d}^3\mathbf{p}_1\mathrm{d}^3\mathbf{p}_2\ldots\mathrm{d}^3\mathbf{p}_{n-2}\mathrm{d}^2\hat{\mathbf{p}}} &\propto \delta(p-p_0)\frac{p\mathrm{d}p}{E'_{n-1}+E'_n}\prod_{i=1}^{n-2}\frac{1}{E_i}\\
&= \left.\frac{p}{E'_{n-1}+E'_n}\prod_{i=1}^{n-2}\frac{1}{E_i}\right|_{p=p_0},
\end{aligned}$$

with p_0 being the value of $p \geq 0$ for which energy is conserved.

p_0 is the root of the delta function, so

$$E'_* = E'_{n-1} + E'_n = \sqrt{p_0^2 + m_{n-1}^2} + \sqrt{p_0^2 + m_n^2},$$

which is of the form

$$\begin{array}{rll} & Q = \sqrt{R} + \sqrt{S} & \\ \Rightarrow & Q - \sqrt{R} = \sqrt{S} & \\ \Rightarrow & Q^2 - 2Q\sqrt{R} + R = S & (*) \\ \Rightarrow & Q^2 + R - S = 2Q\sqrt{R} & \\ \Rightarrow & (Q^2 + R - S)^2 = 4Q^2R. & (*) \end{array}$$

This may be used on the expression now, but note that in the squaring operations marked (*) the implication goes only one way, so any value of p_0 derived using the following must be checked back to make sure the squaring operation did not equate things that originally had opposite signs! The squared-up expression is

$$({E'_*}^2 + m_{n-1}^2 - m_n^2)^2 = 4{E'_*}^2(p_0^2 + m_{n-1}^2),$$

giving the root explicitly as

$$p_0 = \sqrt{\frac{({E'_*}^2 + m_{n-1}^2 - m_n^2)^2}{4{E'_*}^2} - m_{n-1}^2}.$$

3 The Primal Distribution

The previous section gave the probability distribution with respect to an element of the momenta $\mathbf{p}_1 \ldots \mathbf{p}_{n-2}$ and the unit direction $\hat{\mathbf{p}}$ governing the last two particles. Thus if a uniform distribution is used for the $n-2$ momenta and an isotropic one for $\hat{\mathbf{p}}$ in the primal distribution, the $(3n-4)$-dimensional probability density will be exactly proportional to the selection probability.

Generating an isotropic $\hat{\mathbf{p}}$ is easy; what is not so obvious is how to choose a 'domain' in which the $\mathbf{p}_i$ are uniformly spread. As $|\mathbf{p}_i| \leq E_i \leq E = m_0$ (i.e. the mass of the parent particle), there is at least a crude limit on how large the momenta can be. Here note that many of these choices for $\mathbf{p}_1 \ldots \mathbf{p}_{n-2}$ miss the manifold entirely, a fact which becomes apparent when attempting to enforce conservation of energy leads to particles with negative energy, or taking square-roots of negative quantities. In these cases, the sample should be rejected, so in effect there are two levels of rejection sampling, one to get onto the manifold and a second to adjust the density. The first level does not affect the relative proportions between any of the probabilities within the manifold, so has no adverse consequences other than wasted CPU time.

The final ingredient for a workable implementation of rejection sampling is an upper bound on the probability density derived in the previous section, so that it may be divided out to give selection probabilities of no more than 1. By noting that $E_i \geq m_i \Rightarrow 1/E_i \leq 1/m_i$, one obtains

$$\frac{p}{E'_{n-1} + E'_n} \prod_{i=1}^{n-2} \frac{1}{E_i} \leq \frac{p}{E'_{n-1} + E'_n} \prod_{i=1}^{n-2} \frac{1}{m_i},$$

leaving only the front term to deal with. Remembering that $|\mathbf{p}'_{n-1}| = |\mathbf{p}'_n| = p$, it follows that $E'_{n-1} \geq p$ and $E'_n \geq p$ so that the denominator is at least $2p$, resulting in

$$\frac{p}{E'_{n-1} + E'_n} \prod_{i=1}^{n-2} \frac{1}{E_i} \leq \frac{1}{2} \prod_{i=1}^{n-2} \frac{1}{m_i}.$$

The probability to use is therefore

$$\mathcal{P}_1 = \frac{2p}{E'_{n-1} + E'_n} \prod_{i=1}^{n-2} \frac{m_i}{E_i}.$$

This is workable for several cases, but when there are particles of small mass (such as electrons or neutrinos) in the decay products, the ratio $m_i/E_i = 1/\gamma_i$ becomes very small when these acquire a significant fraction of the mass-energy of the heavier parent particle. The formula does not select any samples at all when photons are produced! The problem arises from the choice of a uniform momentum distribution for each of the first $n-2$ particles not being a very good 'fit' to the $1/E_i$ in the probability density, particularly when the energy is dominated by momentum rather than rest mass.

With this in mind, the bound can be modified to use the fact that $E_i = \sqrt{\mathbf{p}_i^2 + m_i^2} \geq \max\{|\mathbf{p}_i|, m_i\}$ so $1/E_i \leq \min\{1/|\mathbf{p}_i|, 1/m_i\}$. However, this is no longer a constant with respect to $\mathbf{p}_i$, so to have it divided out, a primal distribution must be generated with $F(\mathbf{p}_i) \propto \min\{1/|\mathbf{p}_i|, 1/m_i\}$. This is uniform for $|\mathbf{p}_i| \leq m_i$, so retaining the previous cutoff of $|\mathbf{p}_i| \leq m_0$, the ratio of the uniform part to the outer part is

$$\frac{4}{3}\pi m_i^3 \frac{1}{m_i} : \int_{m_i}^{m_0} 4\pi |\mathbf{p}_i|^2 \frac{\mathrm{d}|\mathbf{p}_i|}{|\mathbf{p}_i|} = \frac{m_i^2}{3} : \frac{m_0^2 - m_i^2}{2}.$$

Since $2|\mathbf{p}_i|\mathrm{d}|\mathbf{p}_i| = \mathrm{d}|\mathbf{p}_i|^2$, the outer part corresponds to choosing $|\mathbf{p}_i|^2$ uniformly on $[m_i^2, m_0^2]$. Generating events from these two parts in the ratio above, for each i up to $n-2$, gives the required primal distribution.

Now dividing out $\min\{1/|\mathbf{p}_i|, 1/m_i\}$ instead of just $1/m_i$ gives the new selection probability as

$$\mathcal{P}_2 = \frac{2p}{E'_{n-1} + E'_n} \prod_{i=1}^{n-2} \frac{\max\{|\mathbf{p}_i|, m_i\}}{E_i},$$

which is clearly better conditioned, as terms in the product part never fall below $1/\sqrt{2}$.

4 Results and Efficiency

The two versions of the algorithm described above were implemented and tested on a variety of decay channels for the kaon: the results are shown in table 4. Note that for 3-body and higher decays, there is sometimes a choice of which particles to associate the 'special' momenta $\mathbf{p}_{n-1}$ and $\mathbf{p}_n$ with. This affects the speed of the algorithms, particularly the first version, so these possibilities are broken down individually in the cases where the particles have different mass.

A standard plot that can be used to test the correctness of the routine is the *Dalitz plot* of m_{12}^2 against m_{23}^2, where $m_{ij}^2 = (m_i + m_j)^2 - |\mathbf{p}_i + \mathbf{p}_j|^2$. The correct phase-space distribution ought to be uniformly spread on the allowed region when projected into this plane. Such a plot is shown in figure 4 generated from many runs of the $\mathcal{P}_2$-based version of the algorithm.

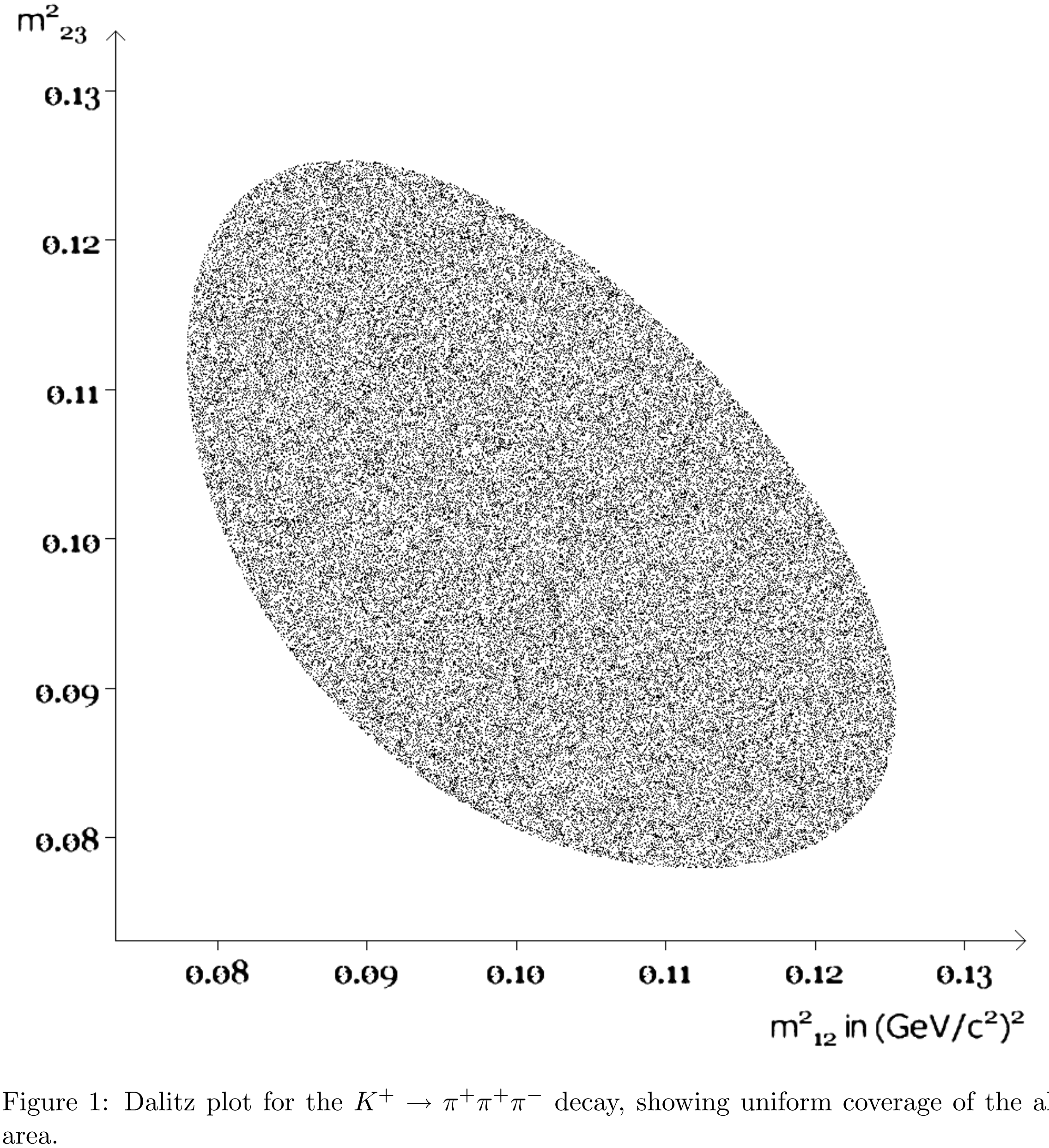

Figure 1: Dalitz plot for the $K^+ \to \pi^+\pi^+\pi^-$ decay, showing uniform coverage of the allowed area.

Table 1: Mean number of primal samples required to generate one valid decay. Figures are an average over 10000 successful decays except where another number of decays is indicated in parentheses. 'inf' denotes that it was infeasible to run the algorithm, as no valid decays were generated after 10^7 samples.

Decay (special pair)	Using $\mathcal{P}_1$	Using $\mathcal{P}_2$
$K^+ \to e^+\nu_e$	1	1
$K^+ \to \mu^+\nu_\mu$	1.0483	1.0485
$K^+ \to \pi^0 e^+\nu_e$		
(π^0, e^+)	inf	6.7499
(π^0, ν_e)	4123.71 (2425)	6.5441
(e^+, ν_e)	15.7419	6.4251
$K^+ \to \pi^+\pi^0\pi^0$		
(π^+, π^0)	151.285	60.5826
(π^0, π^0)	146.47	61.5364
$K^+ \to \pi^+\pi^+\pi^-$	184.181	75.1907
$K^+ \to \pi^+\pi^- e^+\nu_e$		
(π^+, π^-)	inf	666.969
(π, e^+)	inf	652.936
(π, ν_e)	8×10^5 (25)	654.098
(e^+, ν_e)	3746.72 (2669)	636.432