Draft v0.2
08 February, 2006

# LIKELIHOOD IMPLEMENTATION IN TOY MC

## 1 Likelihood Definitions

The likelihood used in the fit is defined as the product of event probability density functions (PDF's).

$$\begin{aligned}\mathcal{L} &= \prod_{e=evts} \mathcal{P}(\vec{x}_e\,;\,\vec{p}) \\ &= \prod_{e} [\theta(+T_e)\mathcal{P}^{unm}(\vec{x}_e\,;\,\vec{p}) + \theta(-T_e)\mathcal{P}^{mix}(\vec{x}_e\,;\,\vec{p})] \qquad (1)\end{aligned}$$

where $\theta(x) = 1$ for $x > 0$ and 0 for $x < 0$, $T$ is the event tag, +1 for events tagged as unmixed and −1 for those tagged as mixed, $\vec{x_e}, \vec{p}$ are vectors of event observables and likelihood function parameters respectively.

The total PDF for an event is the sum of PDFs for each possible source contributing to the data sample weighted by their fractions, $f_a$, in the total tagged sample (all events tagged as either unmixed or mixed):

$$\mathcal{P}^{tag}(\vec{x}\,;\,\vec{p}) = \sum_{a=src} f_a \mathcal{P}_a^{tag}(\vec{x_a}\,;\,\vec{p}) \qquad (2)$$

As we will see below, the values of the event observables may depend on the assumed source.

The total PDF is normalized to unity:

$$\int_{-\infty}^{+\infty} d^n\vec{x}\,\mathcal{P}(\vec{x}\,;\,\vec{p}) = 1 \qquad (3)$$

Details of the implementation of this normalization condition can be found in section 6.

Since there are several observables for each event, which may be correlated, the PDF ($\mathcal{P}$) is given above is the *joint* PDF of all the variables. The *Probability* section of the PDG Review [1] describes how to relate this joint PDF to the conditional PDF that is easy to implement mathematically. In the following we will assume that the first element of the observables vector, $\vec{x}$, corresponds to the "main" variable sensitive to mixing – *e.g.* the VPDL or $L_{xy}$. The joint PDF can then be written as:

$$\mathcal{P}_a(x_1, x_2, \cdots, x_n) = \mathcal{P}_a(x_1 \mid x_2, \cdots, x_n)\mathcal{P}_a(x_2)\cdots\mathcal{P}_a(x_n) \qquad (4)$$

where

- $\mathcal{P}_a(x_1, x_2, \cdots, x_n)$ is the probability of *jointly* observing $x_1, x_2, \cdots, x_n$ in an event for source $a$.
Normalized such that $\int dx_1 dx_2 \cdots dx_n \, \mathcal{P}_a(x_1, x_2, \cdots, x_n) = 1$

- $\mathcal{P}_a(x_1 \mid x_2, \cdots, x_n)$ is the probability of observing $x_1$ *given* $x_2, \cdots, x_n$ for source $a$.
Normalized such that $\int dx_1 \, \mathcal{P}_a(x_1 \mid x_2, \cdots, x_n) = 1$

- $\mathcal{P}_a(x_i)$ is the probability density of variable $x_i$ for source $a$.
Normalized such that $\int dx_i \, \mathcal{P}_a(x_i) = 1$

In the above we have assumed that the "nuisance" variables $x_2, \cdots, x_n$ are uncorrelated, howver, this can be avoided by continuing to construct conditional PDFs for the variables that are left.

## 2 Fit Results Histograms

To check the quality of the fit, it is useful to compare the distributions of data and the prediction of the fit vs. the main variable for samples tagged as "unmixed" and "mixed". This is a far more sensitive indicator of problems with the fit than the result for $\Delta m_s$.

To construct these histograms, the unmixed and mixed PDFs, at the best-fit parameters, $\vec{p}_0$, must be integrated over the secondary variables.

$$\begin{aligned}
\mathcal{P}^{tag}(x_1 \,;\, \vec{p}_0) &= \int dx_2 \cdots dx_n \, \mathcal{P}^{tag}(\vec{x} \,;\, \vec{p}) \\
&= \sum_a f_a \int dx_2 \cdots dx_n \, \mathcal{P}_a^{tag}(x_1 \mid x_2, \cdots, x_n \,;\, \vec{p}_0) \\
&\quad \times \mathcal{P}_a(x_2) \cdots \mathcal{P}_a(x_n) \qquad (5)
\end{aligned}$$

Integrals are approximated by sums in the code.

The histogram, $\mathcal{H}^{tag}(x_1)$, is then

$$\mathcal{H}^{tag}(x_1) = \int_{bin} dx_1 \mathcal{P}^{tag}(x_1 \,;\, \vec{p}_0) \qquad (6)$$

which can be approximated by the value at the bin center for small enough bins.

## 3 Variables Used in the Fit

Various observable sets can be used in the fit. The two sets implemented so far are the following.

### 3.1 Decay Length Based

**Main Variable:**

- $L_{xy}$: the 2D distance from the primary to the secondary vertex.

**Secondary Variables:**

- $P_t^m$: the measured $(D_s + \ell)$ transverse momentum in the event.
- $\sigma_L$: the measured error on $L_{xy}$.
- $T$: the mixing tag – +1=unmixed, -1=mixed
  (note: in the code the convention 0=unmixed, 1=mixed is often used)
- $\mathcal{D}$: the dilution

### 3.2 VPDL Based

**Main Variable:**

- $X_a$: the VPDL (depends on the source assumed – see below)

**Secondary Variables:**

- $\sigma_{xa}$: the measured error on $X_a$.
- $T$: the mixing tag – +1=unmixed, -1=mixed
  (note: in the code the convention 0=unmixed, 1=mixed is often used)
- $\mathcal{D}$: the dilution

### 3.3 Defining the VPDL Correctly

A generic VPDL and its error can be defined for any event as:

$$\begin{aligned} X &= M(B_s)\frac{\vec{L}_{xy} \cdot \vec{P}_t^m}{P_t^{m2}} \sim M(B_s)\frac{L_{xy}}{P_t^m} \\ \sigma_x &\sim M(B_s)\frac{\sigma_L}{P_t^m} \end{aligned} \tag{7}$$

However, this does not take into account the different masses of $B$-mesons. When used as a variable in the PDF for source $a$, the vpdl should be modified to:

$$\begin{aligned} X_a &= X\frac{M(B_a)}{M(B_s)} \\ \sigma_{xa} &= \sigma_x\frac{M(B_a)}{M(B_s)} \end{aligned} \tag{8}$$

unless a mass is not defined for this source ($c\bar{c}$, combinatoric background, etc.) in which case the original VPDL variables are used.

## 4 Likelihood Implementation

The likelihood as implemented in the code consists of several components.

1. Convolution over the $K$-factor.

$$K = \frac{P_t^m}{P_t(B_s)} \tag{9}$$

   Note: $K$=1 for $c\bar{c}$, combinatoric background, etc.

2. Convolution over the resolution function, $\mathcal{R}(x^t - x^m; \sigma_x)$, assumed to be a gaussian with width given by the event error, where $x^{t,m}$ are the true and measured values of the VPDL or $L_{xy}$.

3. Convolution over the efficiency function, $\varepsilon$, as a function of $x^m$.

4. The efficiency function normalization, $A(\vec{p}, K, x_2, \cdots, x_n)$, which depends on the values of the fit parameters. It is defined by the condition:

$$\begin{aligned} A_a(\vec{p}, K, x_2, \cdots, x_n) &= \int dx^m \varepsilon_a(x^m) \int dx^t \mathcal{R}(x^t - x^m; \sigma_x) \\ &\times \left[F_a^{unm}(x^t, \cdots) + F_a^{mix}(x^t, \cdots)\right] \end{aligned} \tag{10}$$

   as discussed in section 6.

5. The physics function, $F_a^{tag}(x^t \mid x_2, \cdots, x_n \,;\, \vec{p})$

6. Histograms for the secondary variables representing their PDFs – $\mathcal{H}_a(v) \sim \mathcal{P}_a(v)$. These are normalized such that: $\sum_{i=bins} \mathcal{H}_a(v_i) = 1$

Convolution integrals over the resolution and efficiency functions are calculated analytically (see section 7 for details), while convolution over the $K$-factor is approximated as a sum over the bins of the $K$-factor histogram.

## 4.1 Decay Length Based PDF

The joint PDF for each source using the decay length based variable set is given by:

$$\mathcal{P}_a^T(\vec{x}\,;\,\vec{p}) = \mathcal{P}_a^T(L_{xy} \mid P_t^m, \sigma_L, T, \mathcal{D}\,;\,\vec{p})\mathcal{H}_a(P_t^m)\,\mathcal{H}_a(\mathcal{D})\,(\mathcal{H}(\sigma_L)) \tag{11}$$

Note that:

- The tag PDF's, $\mathcal{H}_a(T)$ are included in the definition of the unmixed and mixed physics functions as discussed in section 5.
- $\mathcal{H}(\sigma_L)$, the distribution of decay length errors is assumed to be independent of source. It thus factors out of the sum of the individual source PDFs and does not affect the fit. It is only included when plots are made of the predicted number of events vs. $L_{xy}$ based on the fit results.

The conditional PDF is:

$$\begin{aligned} \mathcal{P}_a^T(L_{xy}^m \mid P_t^m, \sigma_L, T, \mathcal{D}\,;\,\vec{p}) &= \sum_k \Delta K \mathcal{H}_a(K_k)\,\frac{\varepsilon_a(L^m)}{A_a(\vec{p}, K, \cdots)} \\ &\times \int_0^\infty dL^t\,\mathcal{R}(L^t - L^m; \sigma_L) \\ &\times F_a^T(L^t \mid K, P_t^m, \sigma_L, \mathcal{D}\,;\,\vec{p}) \end{aligned} \tag{12}$$

### 4.2 VPDL Based PDF

The joint PDF for each source using the VPDL based variable set is given by:

$$\mathcal{P}_a^T(\vec{x}_a\,;\,\vec{p}) = \mathcal{P}_a^T(X_a \mid \sigma_{xa}, T, \mathcal{D}\,;\,\vec{p})\mathcal{H}_a(\sigma_{xa})\,\mathcal{H}_a(\mathcal{D}) \tag{13}$$

Note that:

- $\mathcal{H}_a(\sigma_{xa})$, the distribution of VPDL errors for source $a$, is required to be included in the likelihood definition used in the fit since the width of each source's gaussian distribution will be different.

The conditional PDF is:

$$\begin{aligned}
\mathcal{P}_a^T(X_a^m \mid \sigma_{xa}, T, \mathcal{D}\,;\,\vec{p}) &= \sum_k \Delta K \mathcal{H}_a(K_k)\;\frac{\varepsilon_a(X_a^m)}{A_a(\vec{p}, K, \cdots)} \\
&\times \int_0^\infty dX_a^t\,\mathcal{R}(X_a^t - X_a^m; \sigma_{xa}) \\
&\times F_a^T(X_a^t \mid K, \sigma_{xa}, \mathcal{D}\,;\,\vec{p})
\end{aligned} \tag{14}$$

### 4.3 PDFs for "Zero"-Lifetime Backgrounds

The PDFs for backgrounds arising from sources with true decay lengths that are small compared to the resolution – $c\bar{c}$, combinatoric backgrounds, etc. – are calculated differently than those for other sources. For these backgrounds, a $K$-factor cannot be computed, because there is no true $P_t(B)$ or $M(B)$. Thus the following definitions are used:

$$\begin{aligned}
K &= 1 \\
X_{cc} &= M(B_s)\frac{L_{xy}}{P_t^m} \\
\sigma_{xcc} &= M(B_s)\frac{\sigma_L}{P_t^m}
\end{aligned} \tag{15}$$

Convolution over the $K$-factor is not necessary and, because, the physics function is approximately a delta-function, $\delta(x^t = 0)$, convolution over the resolution function can also be avoided. (Note: even if this is not the case, the physics function for the 0-lifetime case can be constructed to include effects of resolution.) With the 0-lifetime physics function re-defined to include effects of resolution, the PDF for this source can be written as:

**Decay Length Based**

$$\mathcal{P}_{cc}(L_{xy}^m \mid P_t^m, \sigma_L, T, \mathcal{D}\,;\,\vec{p}) = \frac{\varepsilon_{cc}(L^m)}{A_{cc}(\vec{p}, K, \cdots)} F_{cc}^T(L^m \mid P_t^m, \sigma_L, \mathcal{D}\,;\,\vec{p}) \tag{16}$$

**VPDL Based**

$$\mathcal{P}_{cc}(X^m \mid \sigma_x, T, \mathcal{D}\,;\,\vec{p}) = \frac{\varepsilon_{cc}(X^m)}{A_{cc}(\vec{p}, K, \cdots)} F_{cc}^T(X^m \mid \sigma_x, \mathcal{D}\,;\,\vec{p}) \tag{17}$$

# 5 Physics Functions

The physics functions used in the code for various source classes are defined below. All use proper time, $t_0$, defined as:

$$\begin{aligned} ct_0 &= M(B_a)\frac{L_{xy}}{P_t^m}K \quad \text{(decay length based)} \\ &= X_a K \quad \text{(VPDL based)} \end{aligned} \tag{18}$$

The functions are also labeled by their tag, $T$:

$$\begin{aligned} T &= +1 \quad \text{unmixed (unm)} \\ &= -1 \quad \text{mixed (mix)} \end{aligned} \tag{19}$$

They are normalized such that:

$$\int dx_1 \, [F_a^{unm}(x_1 \mid x_2, \cdots, x_n \, ; \, \vec{p}) + F_a^{mix}(x_1 \mid x_2, \cdots, x_n \, ; \, \vec{p})] = 1 \tag{20}$$

This implies that the individual tagged functions, $F_a^{unm}$ and $F_a^{mix}$ include the tagging PDF instead of being normalized to unity as is usual with PDFs:

$$\begin{aligned} \int dx_1 \, F_a^{mix} &= \chi_a \\ \int dx_1 \, F_a^{mix} &= 1 - \chi_a \end{aligned} \tag{21}$$

where $\chi_a$ is the intrinsic probability that this source appears in the mixed sample (*i.e.* the time-integrated mixing probability for those sources that mix). Expressing the physics functions in a more "standard" form would then give:

$$F_a^T = \mathcal{H}_a(T)\mathcal{F}_a^T \tag{22}$$

where the two functions $\mathcal{F}_a^T$ are separately normalized to unity.

## 5.1 Mixing Sources

$$\begin{aligned} F_a^T &= \frac{N}{2}\exp(-t_0/\tau_a)\,[1 + T\mathcal{D}\cos(\Delta m_a t_0)] \\ N &= \frac{M_a K}{c\tau_a P_t^m} \quad \text{(decay length)} \\ N &= \frac{K}{c\tau_a} \quad \text{(VPDL)} \end{aligned} \tag{23}$$

This gives the proper normalization for each tag separately. As an example consider the mixed sample for VPDL based variables with $\mathcal{D} = 1$.

$$\begin{aligned} \int_0^\infty dX \, F_a^{mix} &= \frac{1}{2} - \frac{K}{2c\tau_a}\int_0^\infty dX \, \exp\left(-\frac{KX}{c\tau_a}\right)\cos\left(\frac{\Delta m_a K X}{c}\right) \\ &= \frac{1}{2}\left[1 - \frac{1}{1 + (\Delta m_a/\tau_a)^2}\right] = \chi_a \end{aligned} \tag{24}$$

### 5.2 Long-Lifetime (non-mixing) Sources

$$
\begin{aligned}
F_a^T &= N\frac{[1+T\mathcal{D}(1-2P_{mix})]}{2}\exp(-t_0/\tau_a) \\
N &= \frac{M_a K}{c\tau_a P_t^m} \ \ \text{(decay length)} \\
N &= \frac{K}{c\tau_a} \ \ \text{(VPDL)}
\end{aligned} \tag{25}
$$

where $P_{mix} \equiv \mathcal{P}(T = mix)$ is the probability of the event to appear in the mixed sample. In this case we usually have $\mathcal{P}(T = unm) = 1$ and $\mathcal{P}(T = mix) = 0$. Then, the probability that an event shows up in the mixed sample is just the mis-tag probability: $(1 - \mathcal{D})/2$.

### 5.3 Zero-Lifetime (non-mixing) Sources

$$
\begin{aligned}
F_{cc}^T &= \frac{[1+T\mathcal{D}(1-2P_{mix})]}{2\sqrt{2\pi}\sigma_L}\exp(-\frac{L_{xy}^2}{2\sigma_L^2}) \ \ \text{(decay length)} \\
&= \frac{[1+T\mathcal{D}(1-2P_{mix})]}{2\sqrt{2\pi}\sigma_L}\exp(-\frac{X^2}{2\sigma_X^2}) \ \ \text{(VPDL)}
\end{aligned} \tag{26}
$$

Note also that the dilution cancels in these functions if $P_{mix} = 0.5$ as is usually the case.

## 6 Details of the Likelihood Normalization

As mentioned in section 1, the total event likelihood is normalized such that its integral over all variables is unity. This requirement gives the efficiency normalization factor.

$$
\begin{aligned}
1 &= \int_{-\infty}^{+\infty} d^n\vec{x}\,\mathcal{P}(\vec{x}\,;\,\vec{p}) \\
&= \int d^n\vec{x}\,\sum_a f_a[\theta(+T)\mathcal{P}_a^{unm}(\vec{x}_a;\vec{p}) + \theta(-T)\mathcal{P}_a^{mix}(\vec{x}_a;\vec{p})] \\
&= \sum_a f_a \int d^{n-1}\vec{x}_a \,\frac{\varepsilon_a(x_a)}{A_a(\vec{p},K,\cdots)} \int dx_a^t \mathcal{R}(x_a^t - x_a;\sigma_x) \\
&\times \ [F_a^{unm}(x^t \mid x_2,\cdots,x_n\,;\,\vec{p}) + F_a^{mix}(x^t \mid x_2,\cdots,x_n\,;\,\vec{p})] \\
&\times \ \mathcal{H}_a(x_2)\cdots\mathcal{H}_a(x_{n-1})
\end{aligned} \tag{27}
$$

where, in the last step, we have integrated over $T = +1,-1$ and used the fact that the physics functions implictly contain the tagging variable PDFs.

This equation is satisfied if the efficiency function normalization, $A_a$, is defined as in eqn 10.

## 7 Convolution Integrals

Integrals to define the efficiency function normalization and to do the convolution over the resolution function are performed analytically to save time in the fits. This results in less flexibility since these integrals must be recalculated by hand for each permutation of physics function, resolution function and efficiency function. However, the savings in time is substantial – a factor of ¿10(?) for the efficiency normalization and another factor of up to 3 for the resolution convolution.

In the following, all integrals are taken from [2].

### 7.1 Efficiency Normalization for Mixing and Long Lifetime Sources

The efficiency normalization integral is:

$$\begin{aligned} A_a(\vec{p}, K, x_2, \cdots, x_n) &= \int_{-\infty}^{+\infty} dx^m\, \varepsilon_a(x^m) \int_0^{+\infty} dx^t\, \mathcal{R}(x^t - x^m; \sigma_x) \\ &\times \; [F_a^{unm}(x^t, \cdots) + F_a^{mix}(x^t, \cdots)] \end{aligned} \tag{28}$$

with the following assumptions about the function components

- Efficiency

$$\varepsilon(x) = p_0 \left\{ 1 - \exp(\frac{x^2}{p_1}) \left[ \sum_{i=2} 5 p_i x^{i-2} \right] \right\} \tag{29}$$

- Resolution Function (single gaussian)

$$\mathcal{R}(x - y; \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[ \frac{-(x-y)^2}{2\sigma^2} \right] \tag{30}$$

- Physics Function (mixing or long lifetime)

$$F^{unm}(y) + F^{mix}(y) = N \exp\left( -\frac{ay}{c\tau} \right) \tag{31}$$

The normalizations are:

$$\begin{array}{ccc} & VPDL & L_{xy} \\ a & K & \dfrac{KM_a}{P_t^m} \\ N & \frac{K}{c\tau} & \dfrac{KM_a}{c\tau P_t^m} \end{array} \tag{32}$$

Calculating the $A$ requires evaluating integrals of the following form:

$$\begin{aligned} \int_0^{+\infty} x^{\nu-1} \exp(-\beta x^2 - \gamma x)\, dx &= (2\beta)^{-\nu/2} \Gamma(\nu) \exp\left( \frac{\gamma^2}{8\beta} \right) D_{-\nu}\left( \frac{\gamma}{\sqrt{2\beta}} \right) \\ & \text{(G\&R eqn 3.462)} \end{aligned} \tag{33}$$

where $\Gamma$ is the Gamma function with $\Gamma(n+1) = n!$ for $n =$ integer, and $D_\nu$ are the parabolic cylinder functions. These are defined for the first few values of $\nu$ as:

$$\begin{aligned} D_{-1}(z) &= e^{z^2/4}\sqrt{\frac{\pi}{2}}\left\{1 - \mathrm{erf}(\tfrac{z}{\sqrt{2}})\right\} \\ D_{-2}(z) &= e^{z^2/4}\sqrt{\frac{\pi}{2}}\left\{\sqrt{\frac{2}{\pi}}e^{-z^2/2} - z\left[1 - \mathrm{erf}(\tfrac{z}{\sqrt{2}})\right]\right\} \\ &\text{(G\&R eqn 9.254)} \end{aligned} \tag{34}$$

where erf, erfc are the error function and the complementary error function, defined as:

$$\begin{aligned} \mathrm{erf}(x) &= \frac{2}{\sqrt{\pi}}\int_0^x e^{-u^2}\,du \\ \mathrm{erfc}(x) &= 1 - \mathrm{erf}(x) \end{aligned} \tag{35}$$

Note that the minus sign in the (G&R eqn 9.254) for $D_{-2}$ is an error.

Other values of $D_\nu$ can be derived from the recursion relation

$$D_{p+1}(z) - zD_p(z) + pD_{p-1}(z) = 0 \tag{36}$$

The integral for eqn 28 is then:

$$\begin{aligned} A &= p_0\left\{1 - \frac{n}{\sqrt{2}\sigma\omega}\sum_{i=0}^{3} I_i \int_0^{+\infty} dy\, y^i \exp\left[-\frac{y^2}{{\sigma_{tot}}^2} - \frac{ay}{c\tau}\right]\right\} \\ &= p_0\left\{1 - \frac{N}{\sqrt{2}\sigma\omega}\sum_{i=0}^{3} I_i J_i\right\} \end{aligned} \tag{37}$$

where:

$$\begin{aligned} N &= \text{physics function normalization} \\ \omega &= \frac{1}{2\sigma^2} + \frac{1}{p_1} \\ {\sigma_{tot}}^2 &= 2\sigma^2 + p_1 \\ h &= \frac{a\sigma_{tot}}{2c\tau} \end{aligned} \tag{38}$$

The $I$ terms are functions of the efficiency parameters:

$$\begin{aligned} I_0 &= p2 + \frac{p_4}{2\omega^2} \\ I_1 &= \frac{1}{2\omega^2\sigma^2}\left(p_3 + \frac{3p_5}{2\omega^2}\right) \\ I_2 &= \frac{p_4}{(2\omega^2\sigma^2)^2} \\ I_3 &= \frac{p_5}{(2\omega^2\sigma^2)^3} \end{aligned} \tag{39}$$

The $J$ integrals are:

$$\begin{aligned}
J_0 &= \frac{\sigma_{tot}}{2}\sqrt{\pi}e^{h^2}\mathrm{erfc}(h) \\
J_1 &= \frac{{\sigma_{tot}}^2}{2}\left[1-\sqrt{\pi}he^{h^2}\mathrm{erfc}(h)\right] \\
J_2 &= \frac{{\sigma_{tot}}^3}{2}\left[-h+\frac{\sqrt{\pi}}{2}(1+2h^2)e^{h^2}\mathrm{erfc}(h)\right] \\
J_3 &= \frac{{\sigma_{tot}}^4}{2}\left[(1+h^2)-\frac{\sqrt{\pi}}{2}h(3+2h^2)e^{h^2}\mathrm{erfc}(h)\right] \qquad (40)
\end{aligned}$$

### 7.2 Efficiency Normalization for Zero-Lifetime Sources

Since there is no Resolution convolution for the zero-lifetime sources, the efficiency normalization integral is simplified:

$$A_{cc}(\vec{p}, K, \cdots) = \int_{-\infty}^{+\infty} dx^m \, \varepsilon_a(x^m)[F_{cc}^{unm}(x^m,\cdots)+F_{cc}^{mix}(x^m,\cdots)] \qquad (41)$$

where the total physics function is:

$$F_{cc}^{unm}+F_{cc}^{mix} = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-x^2}{2\sigma^2}\right) \qquad (42)$$

Using the same form for the efficiency function as above, and the integrals

$$\begin{aligned}
\int_{-\infty}^{+\infty} x^{2n}e^{-ax^2}\,dx &= \frac{1\cdot 3\cdot 5\cdots(2n-1)}{2^n a^n}\sqrt{\frac{\pi}{a}} \\
\int_{-\infty}^{+\infty} x^{2n+1}e^{-ax^2}\,dx &= 0 \qquad (43)
\end{aligned}$$

the efficiency normalization is:

$$A_{cc} = p_0\left\{1-\frac{1}{\sqrt{2}\sigma\omega}\left(p_2+\frac{p_4}{2\omega^2}\right)\right\} \qquad (44)$$

### 7.3 Resolution Convolution for Mixing Sources

Resolution convolution is an integral over $y = x^{true}$ of the physics function times resolution function.

$$\begin{aligned}
I_{res}(x) &= \int_0^{+\infty} dy\,\mathcal{R}(x-y;\sigma)F_a^{tag}(y\mid\cdots;\vec{p}) \\
&= \frac{N}{\sqrt{2\pi\sigma^2}}\int_0^{+\infty} dy\,\exp\left(-\frac{(x-y)^2}{2\sigma^2}\right) \\
&\times\ \exp\left(-\frac{ay}{c\tau}\right)\left[1+T\mathcal{D}\cos\left(\Delta m\frac{ay}{c}\right)\right] \qquad (45)
\end{aligned}$$

It involves doing integrals of the form:

$$\begin{aligned}
\int_0^{+\infty} e^{-\beta x^2-\gamma x}\cos(bx)\,dx &= \frac{1}{4}\sqrt{\frac{\pi}{\beta}}\left\{e^{Z^{*2}}[1-\mathrm{erf}(z^*)]+e^{Z^2}[1-\mathrm{erf}(z)]\right\} \\
&= \frac{1}{2}\sqrt{\frac{\pi}{\beta}}\mathcal{R}e\left\{e^{Z^2}[1-\mathrm{erf}(z)]\right\} \\
&\quad \text{(G\&R eqn 3.897)} \qquad (46)
\end{aligned}$$

where

$$z = \frac{\gamma + ib}{2\sqrt{\beta}} \tag{47}$$

and we have use the property of the error function of an imaginary argument:

$$\text{erf}(z^*) = \text{erf}^*(z) \tag{48}$$

Note that complex arguments are not implemented for the `cmath` version of erf. A set of c++ functions to do this was taken from [3].

The resoultion convolution integral for mixing sources using a single gaussian resolution function is:

$$\begin{aligned} I_{res}(x) &= \frac{N}{4} \exp\left[\frac{1}{2}\left(\frac{a\sigma}{c\tau}\right)^2 - \frac{ax}{c\tau}\right] \left\{ \text{erfc}(g) + T\mathcal{D} \exp\left[-\frac{1}{2}\left(\frac{a\sigma\Delta m}{c}\right)^2\right] \right. \\ &\times \left. \mathcal{R}e\left[e^{iM}\text{erfc}(z)\right]\right\} \end{aligned} \tag{49}$$

where $N, a$ are defined above and

$$\begin{aligned} g &= \frac{1}{\sqrt{2}}\left(\frac{a\sigma}{c\tau} - \frac{x}{\sigma}\right) \\ z &= g + i\frac{a\sigma}{\sqrt{2}c}\Delta m \\ M &= 2\mathcal{R}e(z)\,\mathcal{I}m(z) \end{aligned} \tag{50}$$

### 7.4 Resolution Convolution for Long-Lifetime Sources

For these sources, the resolution convolution integral is:

$$\begin{aligned} I_{res}(x) &= \frac{N}{\sqrt{2\pi\sigma^2}} \int_0^{+\infty} dy\; \exp\left(-\frac{(x-y)^2}{2\sigma^2}\right) \\ &\times \exp\left(-\frac{ay}{c\tau}\right)\left[1 + T\mathcal{D}(1 - 2P_{mix})\right] \end{aligned} \tag{51}$$

where $P_{mix}$ is the probability that the true event (without mis-tagging) appears in the mixed sample.

The integral is:

$$I_{res}(x) = \frac{N}{2}\left[1 + T\mathcal{D}(1 - 2P_{mix})\right] \exp\left[\frac{1}{2}\left(\frac{a\sigma}{c\tau}\right)^2 - \frac{ax}{c\tau}\right] \text{erfc}(g) \tag{52}$$

with $g$ defined as above.

## References


[1] http://pdg.lbl.gov

[2] Gradshteyn and Ryzhik, "Tables of Integrals, Series and Products" (1980)

[3] J. Smith, http://www.octave.org/octave-lists/archive/octave-sources.1998/msg00032.html
Note: a bug was found in the function cerf_continued_fraction and fixed in the code used in the fits.