Comments on Art Snyder's "CP Fitting"

R. Cahn – November 22, 1999

Suppose there are tagging categories labeled by a with true dilutions $\overline{D}_a$. The true distributions are described by

$$f_a \frac{\Gamma}{2} e^{-\Gamma t}(1 \pm \overline{D}_a \overline{A} \sin \Delta m t)$$

where f_a is the fraction of events that fall in category a and $\overline{A}$ is the true value of the asymmetry.

The dilutions are the output of some algorithm and their values may be known only to some approximation. We ask what we will find using some assumed values for the dilutions using maximum likelihood to determine A. The log likelihood is a sum over the tagging categories (we ignore here the fluctuations in the number of events falling into each category). We regard the assumed values D_a as fixed and only A is to be determined by maximizing the likelihood.

$$\begin{aligned}
\langle \mathcal{L} \rangle &= \sum_a \tfrac{1}{2} f_a \int dt \frac{\Gamma}{2} e^{-\Gamma t} [(1 + \overline{D}_a \overline{A} \sin \Delta m t) \ln(1 + D_a A \sin \Delta m t) + \\
&\quad (1 - \overline{D}_a \overline{A} \sin \Delta m t) \ln(1 - D_a A \sin \Delta m t)]
\end{aligned}$$

The maximum of the log-likelihood is determined by

$$\begin{aligned}
\langle \frac{\partial \mathcal{L}(A)}{\partial A} \rangle = 0 &= \sum_a \tfrac{1}{2} f_a \int dt \Gamma e^{-\Gamma t} D_a \sin \Delta m t \left[\frac{1 + \overline{D}_a \overline{A} \sin \Delta m t}{1 + D_a A \sin \Delta m t} - \frac{1 - \overline{D}_a \overline{A} \sin \Delta m t}{1 - D_a A \sin \Delta m t} \right] \\
&= \sum_a f_a D_a (\overline{D}_a \overline{A} - D_a A) \int dt \Gamma e^{-\Gamma t} \frac{\sin^2 \Delta m t}{1 - D_a^2 A^2 \sin^2 \Delta m t} \\
&= \sum_a f_a D_a (\overline{D}_a \overline{A} - D_a A) I_1(x = \Delta m / \Gamma, \lambda = D_a A)
\end{aligned}$$

where

$$\begin{aligned}
I_1(x, \lambda) &= \int_0^\infty du e^{-u} \frac{\sin^2 xu}{1 - \lambda^2 \sin^2 xu} \\
&\approx \frac{2x^2}{1 + 4x^2} + \lambda^2 \frac{24 x^4}{(1 + 4x^2)(1 + 16x^2)} \cdots
\end{aligned}$$

We see that the correct value of A (i.e. $\overline{A}$) will be found if

$$\overline{D}_a = D_a$$

that is, if we know and use the proper values for the dilutions of each category. In this case we can compute the expected uncertainty in A

$$\sigma^{-2} = -N\langle\frac{\partial^2\mathcal{L}(A)}{\partial A^2}\rangle = N\sum_a f_a\overline{D}_a^2 I_1(x = \Delta m/\Gamma, \lambda = \overline{D}_a A)$$

Now suppose that we lump together some of the categories a in our assumed values D_a. For example, we might imagine that the dilution of μ-tagged events depends also on the momentum of the μ. What happens if we ignore that dependence in the analysis?

Let α enumerate the sub-categories of a. We will find the true solution if

$$\sum_\alpha f_{a,\alpha} D_a(\overline{D}_{a,\alpha} - D_a) = 0$$

where $f_{a,\alpha}$ is the contribution of subcategory α to the full category a. Here D_a doesn't depend on α, by assumption. We see that getting the right answer requires only that

$$\sum_\alpha f_{a,\alpha}\overline{D}_{a,\alpha} = f_a D_a$$

That is, as long as we have the right average values of the dilution, we get the right output value for the asymmetry, A.

What then is the difference between using small categories and large ones? If we break down a into its subcategories we have the replacement

$$\begin{aligned} f_a D_a^2 &\rightarrow \sum_\alpha f_{a,\alpha} D_{a\alpha}^2 \\ &= \sum_\alpha f_{a,\alpha}[(D_{a\alpha} - D_a)^2 + D_a^2] \\ &= f_a D_a^2 + \sum_\alpha f_{a,\alpha}(D_{a\alpha} - D_a)^2 \end{aligned}$$

This increasing sum means that the expected error decreases.

We get the right answer as long as we use the right averages, but we minimize the uncertainty when we use the maximum information about tagging. Of course we have simply recovered the formula for the total dilution as a sum of contributions to D^2.