◇ Monograph Excerpt ◇

Matter Antimatter Fluctuations

search, discovery and analysis of Bs flavor oscillations

Nuno Leonardo

Complete work published as:
Analysis of B_s oscillations at CDF, MIT Thesis (2006)
Matter antimatter fluctuations, Monograph, LAP Lambert (2011)

Analytical likelihood evaluation

The establishment of probability density functions thoroughly describing the data, in the multiple reconstructed decay channels, and implementation of the associated common fitting framework form a keystone of the analysis enterprise. The complexity of the likelihood model was successively built and derived in the chapters composing the monograph. In this section we address one specific aspect further contributing to the robustness of the implementation and operational performance of the analysis machinery. Namely, the development and employment of PDF expressions allowing precise and efficient likelihood evaluations, specifically through analytical integration.

In general, the computation of probability integrals is less time consuming and more precise when these are solved analytically rather than numerically. At the very beginning one needs to evaluate convolution integrals with normal (resolution) distributions, and these are already complex enough, being solved only through the use of the error function. In cases where integrals of such functions need to be evaluated, not only for each iteration step of the fit, but also for each event input, the computation time, as well as the precision of the numerical method, are of concern. Other cases where multiple integrals of such types are involved require that at least part of the expressions be evaluated analytically in order to render the fits feasible in a reasonable amount of time.

.1 Lifetime

The optimization of the computation speed is dependent on the possibility of evaluating analytically probability integrals, namely the PDF normalization to be performed for each input event. The issue is trivial for unbiased proper decay time distributions (5.15). In cases of explicit biases (5.17) the computation is also readily performed (5.18). The issue is not straightforward however for general biasing effects, induced both at trigger or reconstruction stages. There, the possibility of analytical PDF normalization relies upon convenient parameterizations of the proper decay time efficiency function, $\mathcal{E}(t)$, defined in (5.3).

The required computation is indicated in (5.20). In what follows the parameterization motivated in (5.26) is assumed. The normalization integrals may accordingly be expressed as

$$\begin{aligned}\mathcal{N}_n(\alpha) &= a_n\tau^n \cdot \frac{1}{2\tau}\, e^{\frac{\sigma_t^2}{2}\frac{\alpha^2}{\tau^2}} \cdot \int_{\zeta_n}^{+\infty} \frac{t^n}{\tau^n}\, e^{-(\alpha+\frac{\tau}{\tau_n})\frac{t}{\tau}}\, \mathrm{Erfc}\left(-\frac{\tau}{\sqrt{2}\sigma_t}\frac{t}{\tau}+\frac{\sigma_t}{\sqrt{2}}\frac{\alpha}{\tau}\right) dt \\ &= a_n\tau^n \cdot \frac{1}{2}\, e^{c^2} \cdot \int_{\zeta_n/\tau}^{+\infty} x^n\, e^{-bx}\, \mathrm{Erfc}\left(-ax+c\right)\, dx \end{aligned} \tag{1}$$

where a, b and c denote the following dimensionless positive parameters,

$$a = \frac{\tau}{\sqrt{2}\sigma_t}, \qquad b = \alpha + \frac{\tau}{\tau_n}, \qquad c = \frac{\sigma_t}{\sqrt{2}}\frac{\alpha}{\tau}\ .$$

While for fully reconstructed decays the parameter α is identical to unity, for partially reconstructed proper times it corresponds to the κ-factor, $\alpha = \kappa$. In the latter case, an additional integration over the k-factor parameter is to be performed, $\int_\kappa \ldots \kappa\mathcal{F}(\kappa)d\kappa$. While an analytical implementation of this integration is in principle achievable using identical techniques, it has not been found necessary.

The integration result, for the prototype parameterization

$$\mathcal{E}(t) \;=\; (a_0 + a_1 t + a_2 t^2)\cdot e^{-\frac{t}{\tau_0}} \cdot \theta(t-\zeta)\ , \qquad \alpha = 1 \tag{2}$$

is given by

$$\begin{aligned}\mathcal{N} \;=\; & \frac{1}{2b}\Big\{\left[a_0 + a_1\frac{\tau}{b}(1-b\frac{\sigma_t^2}{\tau\tau_0}) + a_2\frac{\tau^2}{b^2}(2+\frac{\sigma_t^2}{\tau^2}(1-\frac{\tau}{\tau_0})b + \frac{\sigma_t^4}{\tau^2\tau_0^2}b^2)\right] \\ & \cdot\left(2 - \mathrm{Erfc}\left(\frac{-1}{\sqrt{2}}(\frac{\zeta}{\sigma_t}+\frac{\sigma_t}{\tau_0})\right)\right)\cdot e^{\frac{\sigma_t^2}{2\tau_0^2}} \\ & + \left[a_0 + a_1\frac{\tau}{b}(1+b\frac{\zeta}{\tau}) + a_2\frac{\tau^2}{b^2}\left(1+(1+b\frac{\zeta}{\tau})^2\right)\right] \\ & \cdot\mathrm{Erfc}\left(\frac{-1}{\sqrt{2}}(\frac{\zeta}{\sigma_t}-\frac{\sigma_t}{\tau})\right)\cdot e^{-b\frac{\zeta}{\tau}+\frac{\sigma_t^2}{2\tau^2}} \\ & + \left[a_1\tau + a_2\tau^2(\frac{\zeta}{\tau}+\frac{2}{b}-\frac{\sigma_t^2}{\tau\tau_0})\right]\cdot\frac{\sigma_t}{\tau}\sqrt{\frac{2}{\pi}}\cdot e^{-\frac{\zeta}{\tau}\left(\frac{\tau}{\tau_0}+\frac{1}{2}\frac{\zeta}{\tau}\frac{\tau^2}{\sigma_t^2}\right)}\Big\}\ . \end{aligned} \tag{3}$$

Probability integrals

The PDF integrals at hand involve the exponential function, power terms, and the error function, taking the following general form

$$I_n \;=\; \int t^n\, e^{-bt}\, \mathrm{Erfc}(-at+c)\ dt\ . \tag{4}$$

The complementary error function has the following integral representation on the real axis, which is taken as definition,

$$\mathrm{Erfc}(z) \equiv \frac{2}{\sqrt{\pi}} \int_z^{\infty} e^{-u^2} \, du \, .$$

The following useful properties are also noted,

$$\mathrm{Erfc}(0) = 1, \qquad \mathrm{Erfc}(\infty) = 0, \qquad \mathrm{Erfc}(-\infty) = 2, \qquad \frac{\partial}{\partial z}\mathrm{Erfc}(z) = -\frac{2}{\sqrt{\pi}} e^{-z^2} \, ,$$

along with the Taylor series and asymptotic expansions

$$\begin{aligned} \mathrm{Erfc}(x) &= 1 - \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)n!} \qquad &(\text{small } x) \, , \\ \mathrm{Erfc}(x) &= \frac{e^{-x^2}}{x\sqrt{\pi}} \big(1 + \sum_{n=1}^{\infty} \frac{(-1)^n (2n)!}{n!(2x)^{2n}}\big) \qquad &(\text{large } x) \, . \end{aligned}$$

Integrals of the type of (4) can be solved analytically. Calculations have been confirmed using the symbolic program `Mathematica` [99]. Tests of the analytical results against those obtained with numerical integration were performed. Results are expressed below in a form appropriate for *definite integration* of the expressions (4) on a domain of the real axis:

$$\begin{aligned} I_0 &\simeq \frac{1}{b}\left(e^{\frac{b^2}{4a^2} - \frac{bc}{a}} \, \mathrm{Erfc}(-at + c - \frac{b}{2a}) - e^{-bt} \, \mathrm{Erfc}(-at + c) \right) \, , \\ I_1 &\simeq \left(\frac{1}{b^2} - \frac{1}{2a^2} + \frac{c}{ab} \right) e^{\frac{b^2}{4a^2} - \frac{bc}{a}} \, \mathrm{Erfc}(-at + c - \frac{b}{2a}) - \frac{1}{ab\sqrt{\pi}} e^{-(-at+c)^2 - bt} \\ &\quad - \frac{bt+1}{b^2} e^{-bt} \, \mathrm{Erfc}(-at+c) \, , \\ I_2 &\simeq \left(\frac{8a^4 - 2a^2b^2 + b^4}{4a^4b^3} + \frac{2c}{ab^2} - \frac{c}{a^3} + \frac{c^2}{a^2 b} \right) e^{\frac{b^2}{4a^2} - \frac{bc}{a}} \, \mathrm{Erfc}(-at + c - \frac{b}{2a}) \\ &\quad - \left(\frac{t}{ab\sqrt{\pi}} - \frac{b^2 - 4a^2 - 2abc}{2a^3 b^2 \sqrt{\pi}} \right) e^{-(-at+c)^2 - bt} - \frac{(bt+1)^2 + 1}{b^3} e^{-bt} \, \mathrm{Erfc}(-at + c) \, . \end{aligned}$$

.2 Mixing

The description of flavor oscillations involves the introduction of a cosine term in the proper decay time PDF. The latter (7.13) involves the following factors

$$\mathcal{P}_{\exp}(t;\kappa) = \left[e^{-\kappa \frac{t}{\tau}} \theta(t) \right] \otimes G(t;\sigma_t) \cdot \mathcal{E}(t) \, , \tag{5}$$

$$\mathcal{P}_{\cos}(t;\kappa) = \left[e^{-\kappa \frac{t}{\tau}} \theta(t) \cdot \cos(w\kappa t) \right] \otimes G(t;\sigma_t) \cdot \mathcal{E}(t) \, , \tag{6}$$

where w denotes the oscillation frequency, and the parameter κ is provisionally identified with unity (the κ-factor) for fully (partially) reconstructed modes.

The convolution integral in (5) has been already evaluated (5.15) in the context of the lifetime analysis

$$\mathcal{P}_{\rm exp}(t;\kappa) \;=\; \frac{1}{2}e^{-\frac{\kappa}{\tau}(t-\frac{\kappa\sigma_t^2}{2\tau})}\,{\rm Erfc}\left(\frac{\kappa\sigma_t^2-t\tau}{\sqrt{2}\sigma_t\tau}\right)\cdot\mathcal{E}(t)\;. \tag{7}$$

Regarding (6), it can be seen that it formally reduces to (5) upon extension to the complex plane. Indeed, by expanding the cosine in terms of exponential functions of imaginary phase, the following relation holds (with $i=\sqrt{-1}$)

$$\begin{aligned}\mathcal{P}_{\cos}(t;\kappa) &= {\rm Re}\left\{\mathcal{P}_{\rm exp}(t;\alpha)\right\} \qquad \text{with} \quad \alpha=\kappa(1+iw\tau)\;, \qquad (8)\\ &= \frac{1}{2}e^{-\frac{\kappa}{\tau}(t-\frac{\kappa\sigma_t^2}{2\tau})}e^{-\frac{\kappa^2\sigma_t^2w^2}{2}}\cdot{\rm Re}\left\{e^{-iw\kappa(t-\frac{\kappa\sigma_t^2}{\tau})}\,{\rm Erfc}(\frac{\kappa\sigma_t^2-t\tau}{\sqrt{2}\sigma_t\tau}+i\frac{\kappa\sigma_t w}{\sqrt{2}})\right\}\cdot\mathcal{E}(t)\;.\end{aligned}$$

That is, the PDF computations arising in the framework of mixing analyses are accomplished as a complexification of those found for lifetime analyses.

A note on PDF normalization

In Section .1 we have tackled the issue of analytically integrate expressions of the form of (5),

$$\mathcal{N}_{\rm exp}(t;\kappa) \;=\; \int_{-\infty}^{+\infty}\mathcal{P}_{\rm exp}(t;\kappa)\,dt \tag{9}$$

needed for proper decay time PDF normalization in the context of lifetime analyses. As it was addressed in Section 7.3.2, this is also the normalization which is needed in the context of mixing analyses.

While integration of expressions of the type (6),

$$\mathcal{N}_{\cos}(t;\kappa) \;=\; \int_{-\infty}^{+\infty}\mathcal{P}_{\cos}(t;\kappa)\,dt\;, \tag{10}$$

is in general not a requisite for the process of likelihood maximization, they may reveal useful in circumstances such as likelihood projections in tagged subspaces. Such expressions are thus evaluated and implemented in the fitting framework. For the unbiased cases the following is obtained,

$$\mathcal{N}_{\cos}(t) \;=\; \frac{\tau}{1+w^2\tau^2} \qquad\qquad \text{for} \quad \mathcal{E}(t)=1\;.$$

In general, the integration (10) may be obtained by extending the result of (9) to the complex plane, as

$$\mathcal{N}_{\cos}(t;\kappa) \;=\; {\rm Re}\left\{\mathcal{N}_{\rm exp}(t;\alpha)\right\} \qquad \text{with} \quad \alpha=\kappa(1+iw\tau)\;.$$

In particular, for proper decay time biases described by a generic t-efficiency function, parameterized as in (2), the normalization may be obtained correspondingly by evaluating in the complex plane the results obtained in Section .1.

Extension to the complex plane

As pointed out, the likelihood computation for tagged events involves in general the task of evaluating expressions in the complex plane. This requires that a complex class be defined in the fitter framework, in order to handle the basic complex operations.

Additionally, one needs to evaluate the complementary error function of complex argument. For this purpose it is convenient to express the latter as

$$\begin{aligned} \mathrm{Erfc}(z) &= e^{-z^2} W(iz) \ , \\ W(z) &= e^{-z^2} \left[1 + \frac{2i}{\sqrt{\pi}} \int_0^z e^{u^2} \, du \right] \ , \end{aligned}$$

with z denoting a complex number, and $W(z)$ is the complex error function (also called Faddeeva function) which is evaluated with existing numerical algorithms. While this is the implementation adopted, we also mention in passing that the complementary error function may alternatively be expressed in terms of the incomplete gamma function,

$$\begin{aligned} \Gamma(a,x) &\equiv \int_x^\infty u^{a-1} e^{-u} du \\ \mathrm{Erfc}(z) &= \begin{cases} \frac{1}{\sqrt{\pi}} \Gamma(\frac{1}{2}, z^2), & \mathrm{R}e\,(z) > 0 \\ 2 - \frac{1}{\sqrt{\pi}} \Gamma(\frac{1}{2}, z^2), & \mathrm{R}e\,(z) < 0 \end{cases} \end{aligned}$$

for which convenient numerical methods are also available.

.3 Log-likelihood expansion

Once the data samples have been characterized and the associated PDFs of the various input quantities established, the analysis of B_s oscillations contains a final additional step. This involves the introduction of an extra parameter, the amplitude $\mathcal{A}$, and a scanning procedure, which is presented in Section 8.3. The procedure requires that many fits be performed to the amplitude parameter, one for each probe frequency value. Such fits need to be executed considerably more times, in many samples of toy Monte Carlo events, for systematics evaluation.

The likelihood maximization procedure involved in the amplitude scanning can be made more time-effective in a few complementary ways. First of all advantage should be taken from the fact that the amplitude is the only floating parameter, and all remaining fit parameters are fixed. The computation of the various PDFs' terms not depending on the amplitude need to be evaluated once only for each event, and the results cached for use during the maximization process, which is to be repeated for the various probed frequencies.

Furthermore, the likelihood maximization may be achieved in a time-efficient fashion based on log-likelihood expansion and derivation. The event likelihood has a linear dependence on the amplitude parameter. The method can be more simply illustrated expressing (8.3) and its logarithmic series expansion as

$$\begin{aligned} 1+\upsilon\mathcal{A} &= 1+\mathcal{A}\cdot\xi\mathcal{D}\cos(wt)\ , \\ \ln(1+\upsilon\mathcal{A}) &= \upsilon\mathcal{A}-\frac{1}{2}\upsilon^2\mathcal{A}^2+\frac{1}{3}\upsilon^2\mathcal{A}^3-\frac{1}{4}\upsilon^4\mathcal{A}^4+\mathcal{O}(\upsilon^5) \qquad (|\upsilon|<1)\ . \end{aligned}$$

That is, the logarithm of the likelihood is expanded as a polynomial on the amplitude with constant and pre-determined coefficients. Likelihood maximization becomes then reduced to finding the polynomial roots; for lower polynomial degrees, the latter can be achieved analytically.

More generally, as introduced in Section 5.1 and Section 7.1, the likelihood has the following structure,

$$\mathcal{L}=\prod_i\sum_\alpha f_\alpha\mathcal{P}_i^\alpha \qquad \text{with} \qquad \mathcal{P}=L_m\,L_\mathcal{D}\,L_{\sigma_t}\,L_{t,\xi}\ ,$$

where the indices i and α run over the number of events and number of sample components, respectively, and f_α denote the component fractions. For the signal component, the proper decay time likelihood factor (8.6) has the following form

$$L_{t,\xi} \;=\; \frac{p_\epsilon}{1+|\xi|}\frac{1}{\mathcal{N}\tau}\left(\mathcal{P}_{\rm exp}+\mathcal{A}\cdot\xi\mathcal{D}\mathcal{P}_{\rm cos}\right)$$

where $\mathcal{N}$ is evaluated in (3), and $\mathcal{P}_{\rm exp}$ and $\mathcal{P}_{\rm cos}$ are given by (5) and (6), respectively. The likelihood logarithm may accordingly be cast in the form

$$\ln(\mathcal{L}(\mathcal{A})) \;=\; \sum_i\ln\left(\alpha_i+\mathcal{A}\cdot\beta_i(w)\right)=\sum_i\ln\left(1+\mathcal{A}\cdot\upsilon_i(w)\right)+\text{const.}$$

where α, β and υ correspond to combined likelihood factors with no dependence on the amplitude parameter. Making use of the logarithm series expansion, one has

$$\ln(\mathcal{L}(\mathcal{A})) \;=\; \sum_n\left(\sum_i\frac{1}{n}(-1)^{n-1}\upsilon_i^n\right)\cdot\mathcal{A}^n+\text{const.}$$

Alternatively to the standard `Minuit` minimization procedure based on (5.2), the amplitude values which maximize the likelihood may be found as the roots of the polynomial

$$\frac{d\ln(\mathcal{L}(\mathcal{A}))}{d\mathcal{A}} \;=\; \sum_n\left(\sum_i(-1)^{n-1}\upsilon_i^n\right)\cdot\mathcal{A}^{n-1}=0\ .$$

The amplitude uncertainty estimates $\sigma_{\mathcal{A}}$ are obtained from the amplitude values for which the log–likelihood has varied by an amount of 0.5 relative to its found maximum.