Reducing Penguin Pollution

David London

Université de Montréal

May 7, 2013

Talk based on arXiv:1209.1413:

work done in collaboration with B. Bhattacharya and A. Datta.

CP asymmetry (CPA) in $B_s^0(t) \to f$ measures $\mathrm{Im}\left[e^{-i\phi_s}(\bar{A}_f/A_f)\right]$, where ϕ_s is the phase of B_s^0-$\bar{B}_s^0$ mixing and A_f ($\bar{A}_f$) is the amplitude for $B_s^0(\bar{B}_s^0) \to f$. Standard mode: $f = J/\psi\phi$. Assuming that A_f is dominated by diagrams with a single weak phase, the indirect CPA cleanly measures $\sin 2\beta_s$, where $\beta_s = \arg[-(V_{tb}^* V_{ts}/V_{cb}^* V_{cs})]$.

Look more closely at $A(B_s^0 \to J/\psi\phi)$:

$$\begin{aligned} \mathcal{A}_s^{J/\psi\phi} &= V_{cb}^* V_{cs}(C' + P'_{ct} - \frac{2}{3}P'_{EW}) + V_{ub}^* V_{us}(P'_{ut} - \frac{2}{3}P'_{EW}) \\ &\equiv A_1 + e^{i\gamma} A_2 \ . \end{aligned}$$

Second term has different weak phase from the first. Its inclusion in the amplitude will ruin the cleanliness of the measurement of $2\beta_s$: called "penguin pollution" (PP).

Is the PP important? First, $|V_{ub}^* V_{us}/V_{cb}^* V_{cs}| \sim \lambda^2$. Second, matrix elements of P'_{ut}, P'_{EW} much smaller than matrix element of C. Estimate: $|A_2/A_1| = O(10^{-3})$. Conclusion: PP is negligible.

But wait: LHCb has measuremed ϕ_s ($= -2\beta_s$ in SM). They find

$$\phi_s = 0.07 \pm 0.09 \ (\mathrm{stat}) \pm 0.01 \ (\mathrm{syst}) \ \mathrm{rad} \ ,$$

in agreement with the SM. (Still, the errors are large enough that NP cannot be excluded.)

However, small ϕ_s reintroduces the question of the size of the PP. Suppose a future measurement finds $\phi_s = -0.1 \pm 0.01 \ \mathrm{rad}$ (error is statistical only). This suggests NP (in SM, $2\beta_s = 0.036 \pm 0.002$ rad). But perhaps the size of the PP has been underestimated (e.g., difficult-to-calculate hadronic matrix elements). If so, there is really a significant theoretical error associated with this measurement.

How was ϕ_s measured? J/ψ and ϕ are vector mesons $\Longrightarrow$ amplitude has 3 helicities, 1 longitudinal (A_0), 2 transverse ($A_\|$, $A_\perp$). Perform angular analysis of the decay $B_s^0 \to J/\psi(\to \ell^+\ell^-)\phi(\to K^+K^-)$. Key point: one can differentiate the 3 helicities by their contributions to the angular distribution.

Aside: $\phi \to K^+K^- \Longrightarrow \exists$ background from K^+K^- pair with relative angular momentum $l=0$ (S-wave). This constitutes a fourth possible helicity. For simplicity of presentation, I will ignore this. However, (i) LHCb included it in their analysis, (ii) all 4 helicities were included in our analysis in arXiv:1209.1413.

A time-dependent angular analysis is performed. We write

$$\frac{d^4\Gamma(B_s^0 \to J/\psi\phi)}{dt\,d\vec{\Omega}} \propto \sum_{k=1}^{6} h_k(t) f_k(\vec{\Omega}) \ .$$

$\vec{\Omega} = (\theta, \psi, \phi)$ in the transversity basis (the angles are defined w.r.t. the momenta of the J/ψ, ϕ and their decay products).

In the presence of B_s^0-$\bar{B}_s^0$ mixing, the two physical states B_{light} and B_{heavy} have a mass difference (Δm_s) and a width difference ($\Delta\Gamma_s$). The $h_k(t)$ can be written

$$
\begin{aligned}
h_k(t) \quad = \quad & \frac{1}{2} e^{-\Gamma_s t} \left[c_k \cos \Delta m_s t + d_k \sin \Delta m_s t \right. \\
& \qquad \left. + \, a_k \cosh \left(\Delta\Gamma_s/2 \right) t + b_k \sinh \left(\Delta\Gamma_s/2 \right) t \right] \ .
\end{aligned}
$$

The a_k-d_k are functions of A_h, $\bar{A}_h$ and ϕ_s. (In going from the decay to the CP-conjugate decay, $A_h \to \eta_h \bar{A}_h$, where $\bar{A}_h = A_h$, but with weak phases of opposite sign, and $\eta_h = +1$ ($h = 0, \|$) or -1 ($h = \perp$).)

The a_k-d_k can be measured with the time-dependent angular analysis. LHCb explicitly neglect the PP by setting $A_2 = 0$ for all helicities in $A = A_1 + e^{i\gamma} A_2$ (the “1-amplitude method”). In this case, $\bar{A}_h = A_h$. There are 6 unknown parameters: the magnitudes of the A_h (3), the relative strong phases (2), and ϕ_s.

But there are many more observables. For example:

$$
\begin{aligned}
h_1 \quad &: \quad a_1 = 2|A_0|^2 \ , \quad c_1 = 0 \ , \quad d_1 = 2|A_0|^2 \sin\phi_s \ , \\
h_2 \quad &: \quad a_2 = 2|A_\||^2 \ , \quad c_2 = 0 \ , \quad d_2 = 2|A_\||^2 \sin\phi_s \ , \\
h_3 \quad &: \quad a_3 = 2|A_\perp|^2 \ , \quad c_3 = 0 \ , \quad d_3 = 2|A_\perp|^2 \sin\phi_s \ .
\end{aligned}
$$

$\exists$ several different ways of extracting ϕ_s. The problem here is that there is a theoretical error due to the neglect of the PP. We have an estimate of its size, but this estimate is itself uncertain. Can we not do better?

Yes: don't neglect A_2, i.e. allow for $\bar{A}_h \neq A_h$ (the "2-amplitude method"). This inequality can arise due to enhanced PP within the SM and/or to the presence of NP. Now the a_k-d_k are expressed in terms of 12 unknown parameters: the magnitudes of the A_h and $\bar{A}_h$ (6), their relative phases (5), and ϕ_s. Define: $\delta_{ij} \equiv \arg(A_i) - \arg(A_j)$, $\bar{\delta}_{ij} \equiv \arg(\bar{A}_i) - \arg(\bar{A}_j)$, $D_{ij} \equiv \arg(\bar{A}_i) - \arg(A_j)$. Note: only 5 phases are independent, e.g., 2 δ_{ij}'s, 2 $\bar{\delta}_{ij}$'s, and D_{00}.

Now

$$h_1 \;:\; a_1 + c_1 = 2|A_0|^2 \;,\; a_1 - c_1 = 2|\bar{A}_0|^2 \;,\; d_1 = 2|A_0||\bar{A}_0| \sin(\phi_s - D_{00}) \;,$$
$$h_2 \;:\; a_2 + c_2 = 2|A_\||^2 \;,\; a_2 - c_2 = 2|\bar{A}_\||^2 \;,\; d_2 = 2|A_\|||\bar{A}_\|| \sin(\phi_s - D_{\|\|}) \;,$$
$$h_3 \;:\; a_3 + c_3 = 2|A_\perp|^2 \;,\; a_3 - c_3 = 2|\bar{A}_\perp|^2 \;,\; d_3 = 2|A_\perp||\bar{A}_\perp| \sin(\phi_s - D_{\perp\perp}) \;.$$

Once the a_k-d_k are measured, can extract $|A_h|$ and $|\bar{A}_h|$ (see above), as well as the δ_{ij}'s and $\bar{\delta}_{ij}$'s. However, the b_k and d_k terms are all functions of $\phi_s - D_{ij} \Longrightarrow$ cannot separate ϕ_s and the D_{ij} solely from the data. Even in this case, where the PP has been retained, theoretical input is necessary.

Theoretical input: $D_{00} = 0$. Why? (i) Other choices lead to shifts of ϕ_s and the D_{ij} by known quantities; (ii) As we will see, the D_{ij} are all expected to be small; (iii) $D_{00} = 0$ is same as in the 1-amplitude method.

Given $A_0 = A_{1,0} + e^{i\gamma} A_{2,0}$ (the $A_{2,0}$ term represents the PP), $D_{00} = \arg(A_0^* \bar{A}_0) \Longrightarrow$ the theoretical error associated with the assumption $D_{00} = 0$ is $O(|A_{2,0}|/|A_{1,0}|)$.

How is this an improvement? Point: while the 2-amplitude method requires one assumption – $D_{00} = 0$ – the 1-amplitude method in fact requires 6 assumptions:

$$|\bar{A}_h| = |A_h| \ (h = 0, \|, \perp) \ , \ \ \bar{\delta}_{ij} = \delta_{ij} \ (ij = \|0, \perp 0) \ , \ \ D_{00} = 0 \ .$$

But there is a theoretical error associated with each assumption.

That is,

$$E(1) \le E(2) \le ... \le E(6) \ ,$$

where $E(N)$ is the theoretical error on ϕ_s when N assumptions are made. While the theoretical error associated with each assumption is only $O(|A_2/A_1|)$, given that multiple assumptions are involved, $E(6)$ is quite a bit larger than $E(1)$. The theoretical error of the 2-amplitude method could well be an order of magnitude smaller than that of the 1-amplitude method.

Note: one can proceed more slowly than jumping from 6 assumptions to 1 assumption. One can relax one assumption of the 1-amplitude method and add a new unknown parameter. The new fit will give the preferred value of ϕ_s. This can be compared with the value found in the full 1-amplitude method, so this will give $E(6) - E(5)$, where the 5 assumptions of $E(5)$ are those that have been retained.

LHCb has performed an analogous analysis. They take $|\bar{A}_h| \neq |A_h|$, but assume this difference to be helicity-independent. That is, they take $|\bar{A}_h|/|A_h| = \lambda \; \forall h$, and add the unknown parameter λ to the fit. The idea is to see how much λ deviates from 1. They find that this deviation can be at most $\pm 5\%$, and leads to a deviation in ϕ_s of ± 0.02 rad. This may be the typical theoretical error associated with each assumption.

Point: to maximully reduce the effect of PP, it is best to perform the fit with the fewest possible assumptions, i.e., adding as many unknown parameters as possible. If a fit with 11 unknowns (the full 2-amplitude method) is not feasible, one can add assumptions, reduce the number of unknowns, and incur a larger theoretical error in ϕ_s. That is, one has a choice in the type of analysis that is performed.

Finally, I asked the hypothetical question: suppose a future measurement finds $\phi_s = -0.1 \pm 0.01 \ \mathrm{rad}$ (error is statistical only). This disagrees with the SM prediction ($2\beta_s = 0.036 \pm 0.002$ rad) and therefore suggests NP. Can we conclude that NP is indeed present when the theoretical error on ϕ_s is taken into account?

If ± 0.02 rad (LHCb) is a typical theoretical error associated with each assumption, then the total theoretical error of the 1-amplitude method could easily be ± 0.1 rad because many assumptions are made. Thus one cannot conclude that NP is present.

On the other hand, in the 2-amplitude method, the theoretical error is still sufficiently small (± 0.02 rad) that the presence of NP can be confirmed. We therefore see that there are real advantages in the search for NP to employing the 2-amplitude method.