Unbinned model-independent measurements with quantum-correlated D^0 decays

Anton Poluektov

University of Warwick, UK

8 February 2018

Introduction

- Model-independent approach to descibe multubody D decays can be (and is) used in various measurements
  - CKM angle γ from $B \to DK$, $DK\pi$
  - CKM angle β from $B^0 \to Dh^0$, Dhh
  - Charm mixing and CP violation in charm
- Uses the fact that we can obtain both the magnitude and phase difference between D^0 and $\overline{D}^0$ from data (using quantum-correlated $D^0\overline{D}^0$ pairs from CLEO/BESIII)
- Uses **bins** (piecewise uniform function) to approximate varying amplitude; constructed such that even though approximation is rough, physics observables are **unbiased**.

- Here I will show that binned approximation is not the only one possible.
- Use γ measurement in $B \to DK, D \to K^0_{\rm S}\pi^+\pi^-$ as an example, but technique could be extended to multibody B decay fits ($B \to DK\pi$) and time-dependent fits (α, charm mixing).
- Refer to [arXiv:1712.08326] for details

Reminder: γ from $B \to DK$, $D \to K^0_S\pi^+\pi^-$

[GGSZ, 2003; Bondar, 2002]

Information on γ from Dalitz plot analysis of $D \to K^0_S\pi^+\pi^-$ from $B \to DK$.

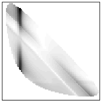
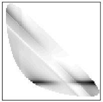
Dalitz plot density: $d\sigma(m^2_+, m^2_-) \sim |A|^2 dm^2_+ dm^2_-$, where $m^2_\pm = m^2_{K_S\pi^\pm}$

Flavor D amplitude: $A_D(m^2_+, m^2_-) \equiv A_D(\mathbf{z})$

CP-conservation in $D \to K^0_S\pi^+\pi^-$ decays: $\overline{A}_D(m^2_+, m^2_-) = A_D(m^2_-, m^2_+)$

Amplitude of $D \to K^0_S\pi^+\pi^-$ from $B^+ \to DK^+$:

$$A_B(m^2_+, m^2_-) = A_D(m^2_+, m^2_-) + r_B e^{i\delta_B + i\gamma} A_D(m^2_-, m^2_+)$$

$=$  $+\, r_B e^{i\delta_B + i\gamma}$ 

Need to know $A_D(m^2_+, m^2_-)$, both amplitude and phase (or, more precisely, phase difference between (m^2_+, m^2_-) and (m^2_-, m^2_+)).

Model-dependent: obtain A_D from $D \to K^0_S\pi^+\pi^-$ fit to the isobar model $\Rightarrow$ model uncertainty

Model-independent: obtain phase difference info from $e^+e^- \to D^0\overline{D}^0$ decays.

Reminder: binned model-independent technique

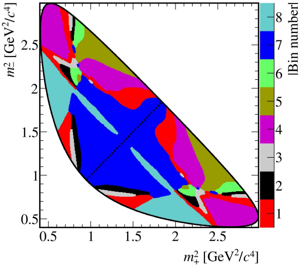
Model-independent: use binned Dalitz plot, deal with bin yields.
Use symmetric binning, $m_+ > m_-$ for $i > 0$, $m_+ < m_-$ for $i < 0$ bins.
Relation between bin yields for $B^+ \to DK^+$ and flavour-specific D.

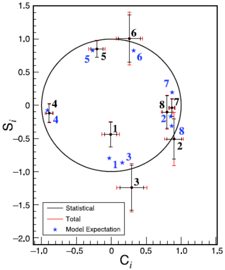
$$N_{\pm i}(B^+) = h_{B^+}[K_i + r_B^2 K_{-i} + 2\sqrt{K_i K_{-i}}(x_+ c_i + y_+ s_i)]$$

(+c.c., which is mosty omitted in this presentation).

$x_\pm = r_B \cos(\delta_B \pm \gamma)$, $y_\pm = r_B \sin(\delta_B \pm \gamma)$ are free parameters.

To reach optimal precision, need bins where phase difference is $\sim$ constant, so the amplitudes add up coherently across bin area.





Model-inspired binning used by CLEO (8×2 bins). [CLEO, PRD82 (2010)]

Phase coefficients c_i, s_i

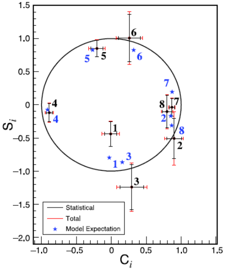
$c_i = \langle \cos\Delta\delta_D \rangle$, $s_i = \langle \sin\Delta\delta_D \rangle$ measured by CLEO (BESIII) in $e^+e^- \to D^0\overline{D}^0$.
Density of correlared $D \to K_S^0\pi^+\pi^-$ Dalitz plots:

$$p_{DD}(m_+^2, m_-^2, {m'_+}^2, {m'_-}^2) \propto |f_D\overline{f}'_D - \overline{f}_D f'_D|^2 \propto$$
$$p_D\overline{p}'_D + p'_D\overline{p}_D - 2\sqrt{p_D\overline{p}'_D p'_D\overline{p}_D}(cc' + ss')$$

After binning:

$$M_{ij} \propto K_i K_{-j} + K_{-i}K_j - 2\sqrt{K_i K_{-j} K_{-i} K_j}(c_i c_j + s_i s_j)$$

which gives c_i, s_i in the fit.



c_i, s_i are aligned around a circle, and their values are well consistent with the calculations from the $D \to K_S^0\pi^+\pi^-$ model.

Do we really need 16 independent parameters to describe an (almost) circle in the phase?

Not really, but then need to go beyond simple binned approximation.

Phase coefficients c_i, s_i

$c_i = \langle \cos\Delta\delta_D \rangle$, $s_i = \langle \sin\Delta\delta_D \rangle$ measured by CLEO (BESIII) in $e^+e^- \to D^0\overline{D}^0$.
Density of correlared $D \to K^0_S\pi^+\pi^-$ Dalitz plots:

$$p_{DD}(m_+^2, m_-^2, m_+'^2, m_-'^2) \propto |f_D\bar{f}'_D - \bar{f}_D f'_D|^2 \propto$$
$$p_D\bar{p}'_D + p'_D\bar{p}_D - 2\sqrt{p_D\bar{p}'_D p'_D\bar{p}_D}(cc' + ss')$$

After binning:

$$M_{ij} \propto K_i K_{-j} + K_{-i}K_j - 2\sqrt{K_i K_{-j} K_{-i} K_j}(c_i c_j + s_i s_j)$$

which gives c_i, s_i in the fit.

c_i, s_i are aligned around a circle, and their values are well consistent with the calculations from the $D \to K^0_S\pi^+\pi^-$ model.

Do we really need 16 independent parameters to describe an (almost) circle in the phase?

Not really, but then need to go beyond simple binned approximation.

Model-independent formalism with weighted integrals

Charm data observables:

$$p_D(\mathbf{z}) = |A_D(\mathbf{z})|^2, \quad \bar{p}_D(\mathbf{z}) = |\overline{A}_D(\mathbf{z})|^2$$

$B^\pm \to DK^\pm$ **data observables**:

$$\bar{p}_B(\mathbf{z}) \propto p_D(\mathbf{z}) + r_B^2 \bar{p}_D(\mathbf{z}) + 2[x_+ C(\mathbf{z}) - y_+ S(\mathbf{z})]$$
$$p_B(\mathbf{z}) \propto \bar{p}_D(\mathbf{z}) + r_B^2 p_D(\mathbf{z}) + 2[x_+ C(\mathbf{z}) + y_+ S(\mathbf{z})]$$

Quantum-correlated $D^0\overline{D}^0$ data observables:

$$p_{DD}(\mathbf{z}_1, \mathbf{z}_2) \propto p_D(\mathbf{z}_1)\bar{p}_D(\mathbf{z}_2) + p_D(\mathbf{z}_2)\bar{p}_D(\mathbf{z}_1) - 2\left[C(\mathbf{z}_1)C(\mathbf{z}_2) + S(\mathbf{z}_1)S(\mathbf{z}_2)\right]$$

Unknowns:

$$C(\mathbf{z}) = \mathrm{Re}\left[A_D^*(\mathbf{z})\overline{A}_D(\mathbf{z})\right], \quad S(\mathbf{z}) = \mathrm{Im}\left[A_D^*(\mathbf{z})\overline{A}_D(\mathbf{z})\right].$$

We want to relate $p_D(\mathbf{z})$, $\bar{p}_B(\mathbf{z})$ and $p_{DD}(\mathbf{z}_1, \mathbf{z}_2)$ and eliminate $C(\mathbf{z}), S(\mathbf{z})$.
We need a way to do it with scattered experimental data.

Model-independent formalism with weighted integrals

Trick: replace all functions $f(\mathbf{z}) \to \int\limits_{\mathcal{D}} f(\mathbf{z}) w_n(\mathbf{z}) d\mathbf{z}$

where $f(\mathbf{z}) = p_D(\mathbf{z}), \bar{p}_B(\mathbf{z}), C(\mathbf{z}), S(\mathbf{z})$.

$w_n(\mathbf{z})$, $1 \leq n \leq N$ is a family of certain weight functions.

Similarly, $p_{DD}(\mathbf{z}_1, \mathbf{z}_2) \to \int\limits_{\mathcal{D}} p_{DD}(\mathbf{z}_1, \mathbf{z}_2) w_m(\mathbf{z}_1) w_n(\mathbf{z}_2) d\mathbf{z}_1 d\mathbf{z}_2$

All the equations will still hold, for any $w_n(\mathbf{z})$.
For scattered data, replace integrals by sums over individual events.

Binned approach is a particular case with

$$w_n(\mathbf{z}) = \begin{cases} 1 \text{ if } \mathbf{z} \in \mathcal{D}_n \\ 0 \text{ otherwise} \end{cases} \quad \text{for bins defined by } \mathcal{D}_n.$$

Altermative approach: **Fourier analysis of the modelled phase difference**

$$w_{2n}(\mathbf{z}) = \sin n\Phi(\mathbf{z}), \quad w_{2n+1} = \cos n\Phi(\mathbf{z})$$

where

$$\Phi(\mathbf{z}) = \arg A_D^{(\mathrm{model})}(\mathbf{z}) - \arg \overline{A}_D^{(\mathrm{model})}(\mathbf{z})$$

Fourier analysis approach

Let's see why Fourier analysis approach is more optimal

We already use model to define optimal binning, the first step is just a continuous generalisation: instead of bins in $\Delta\phi$ we use a continuous variable ϕ. E.g. for $D \to K_S^0\pi^+\pi^-$ Dalitz plot, define

$$\Phi(m_+^2, m_-^2) = \arg A_D^{(\mathrm{model})}(m_+^2, m_-^2) - \arg A_D^{(\mathrm{model})}(m_-^2, m_+^2)$$

Instead of binning in $\Phi(\mathbf{z})$, deal with continuous 1D distribution:

$$p_D(\phi) = \int\limits_{\Phi(\mathcal{D})=\phi} p_D(\mathcal{D})d\mathcal{D}$$

It is just a PDFs of the $\phi = \Phi(\mathcal{D})$ variable for the flavour D sample ($\mathcal{D} \equiv (m_+^2, m_-^2)$).

Define all functions of ϕ similarly ($\bar{p}_B(\phi)$, $p_{DD}(\phi_1, \phi_2)$, $C(\phi)$, $S(\phi)$)

- $C(\phi)$ is even, $S(\phi)$ is odd by construction

Fourier analysis approach

Now we have relation between 1D $p_D(\phi)$ distributions for flavour D and 2D $p_{DD}(\phi_1, \phi_2)$ distributions for correlated $D^0 \overline{D}^0$:

$$p_{DD}(\phi_1, \phi_2) \propto p_D(\phi_1)\bar{p}_D(\phi_2) + \bar{p}_D(\phi_1)p_D(\phi_2) - 2\left[C(\phi_1)C(\phi_2) + S(\phi_1)S(\phi_2)\right]$$

This constrains $S(\phi)$ and $C(\phi)$, which we can apply to $B^+ \to DK^+$:

$$\bar{p}_B(\phi) \propto p_D(\phi) + r_B^2 \bar{p}_D(\phi) + 2[x_+ C(\phi) - y_+ S(\phi)],$$

and extract x and y (and thus γ, after adding B^-).

We still need a way to parametrise functions $C(\phi)$ and $S(\phi)$ to deal with scattered data.

These functions are continuous, periodic, and *resemble* $\cos\phi$ and $\sin\phi$, so Fourier series is a natural parametrisation.

- Calculating Fourier transformation of scattered data is easy.
- Most of the information will already be contained in the leading Fourier term!

Equations for coefficients of Fourier series

So, $p_D(\phi)$ will be parametrised by Fourier series:

$$p_D(\phi) = \frac{a_0^D}{2} + \sum_{n=1}^{M} [a_n^D \cos(n\phi) + b_n^D \sin(n\phi)],$$

similarly for $\bar{p}_B(\phi)$ (coeffs a_n^B, b_n^B),
$C(\phi)$ (coeffs a_n^C, only cosine terms),
$S(\phi)$ (coeffs b_n^S, only sine terms)
$p_{DD}(\phi_1, \phi_2)$ (coeffs a_{nm}^{DD}, b_{nm}^{DD}, c_{nm}^{DD}, and d_{nm}^{DD})

They will be related as

$$\begin{aligned} a_{mn}^{DD} &= 2h_{DD}\left(a_m^D a_n^D - a_m^C a_n^C\right), \\ b_{mn}^{DD} &= c_{mn}^{DD} = 0, \qquad\qquad \text{for } D^0\overline{D}^0 \text{ data} \\ d_{mn}^{DD} &= -2h_{DD}\left(b_m^D b_n^D + b_m^S b_n^S\right). \end{aligned}$$

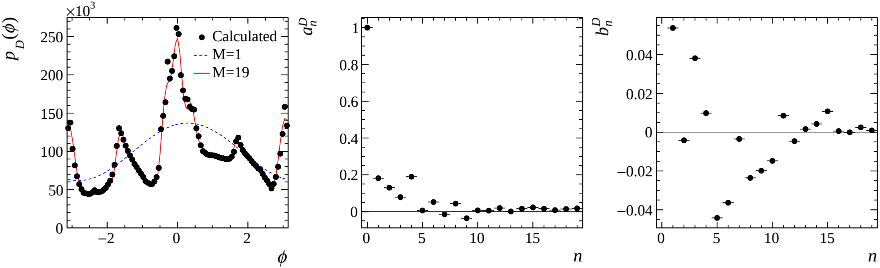
and

$$\begin{aligned} a_n^B &= h_B\left[(1+r_B^2)a_n^D + 2x_- a_n^C\right], \\ b_n^B &= h_B\left[(1-r_B^2)b_n^D + 2y_- b_n^S\right]. \qquad \text{for } B \to DK \text{ data} \end{aligned}$$

System of equations, can be solved for any $M \geq 1$ with maximum likelihood fit

Toy MC results

Large flavour-specific $D \to K^0_S \pi^+ \pi^-$ sample: 10^7 events

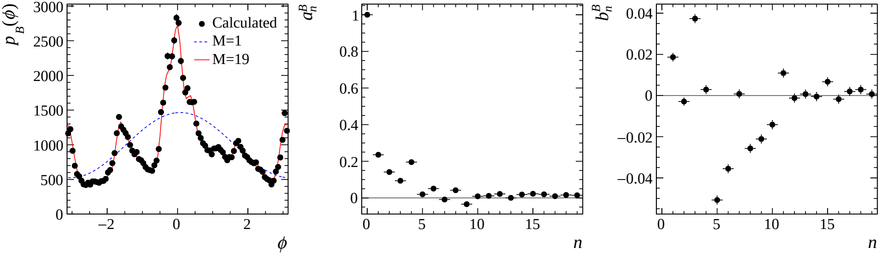


$p_D(\phi)$ density and its spectral coefficients.

Blue line is expansion up to $M = 1$ (first term in Fourier series)

Toy MC results

Large $B^+ \to DK^+$ sample: 10^5 events

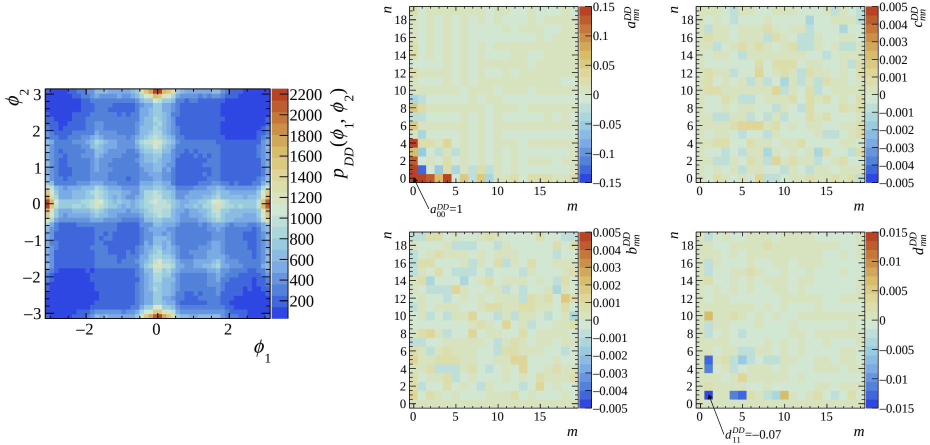


$\bar{p}_B(\phi)$ density and its spectral coefficients.

Blue line is expansion up to $M = 1$ (first term in Fourier series)

Toy MC results

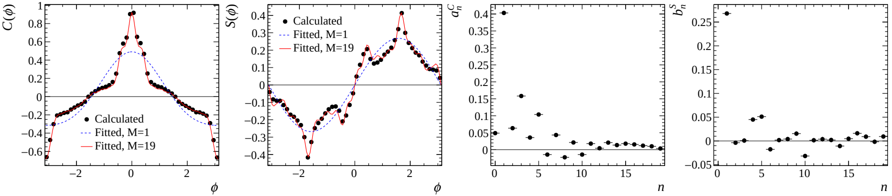
Large quantum-correlated $D\overline{D}$ sample: 10^6 events



$p_{DD}(\phi_1, \phi_2)$ density and its (2D) spectral coefficients

Toy MC results

This allows us to calculate $(a, b)_n^{C,S}$ and reconstruct $C(\phi)$ and $S(\phi)$ (although to measure γ we don't need explicit functions, only coefficients)



Spectral coefficients for $C(\phi)$ and $S(\phi)$ functions and reconstructed functions

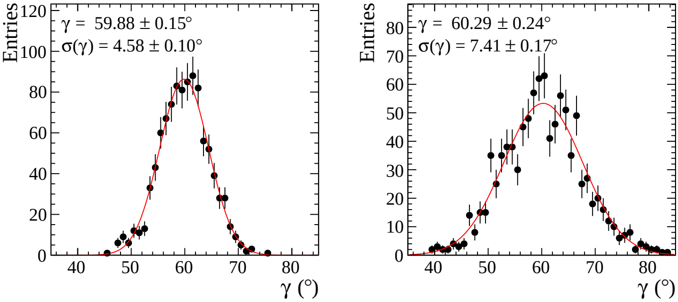
As expected, largest "power" is in the 1st harmonics, $\cos\phi$ term for $C(\phi)$ and $\sin\phi$ term for $S(\phi)$

- With $M = 1$, only 3 free parameters of the D^0 amplitude: a_0^C, a_1^C and b_1^S

Toy MC results

Finally, perform γ fit. 10^4 $B^+ \to DK^+$ decays with $\gamma = 60^\circ$, 1000 toy samples.

Check that the method is $\sim$unbiased for any number of harmonics M, as well as if "wrong" model is used for phase variation $\Phi(m_+^2, m_-^2)$:

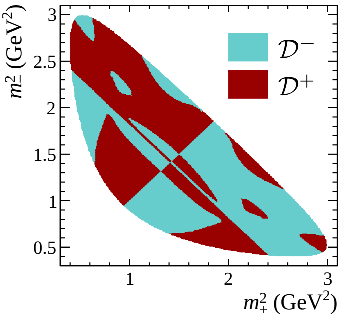


Reduced model: a simplified $D \to K_S^0\pi^+\pi^-$ model with only $K^*(892)$, ρ, ω, $f_0(980)$ + flat non-resonant term, used only for phase difference mapping (*i.e.* different from the generated one).

Further improvement: split Dalitz plot

Using only phase difference information is not optimal, because regions with different $|A_D|$ are treated equally. In the binned approach: "optimal" binning from stochastic optimisation of a certain FoM.

A straightforward solution for Fourier analysis: split Dalitz plot.



Split phase space into two regions, with

$$|A_D^{(\mathrm{model})}(\mathbf{z}) > |\overline{A}_D^{(\mathrm{model})}(\mathbf{z})|| \qquad (\mathcal{D}^+)$$

and

$$|A_D^{(\mathrm{model})}(\mathbf{z}) < |\overline{A}_D^{(\mathrm{model})}(\mathbf{z})|| \qquad (\mathcal{D}^+)$$

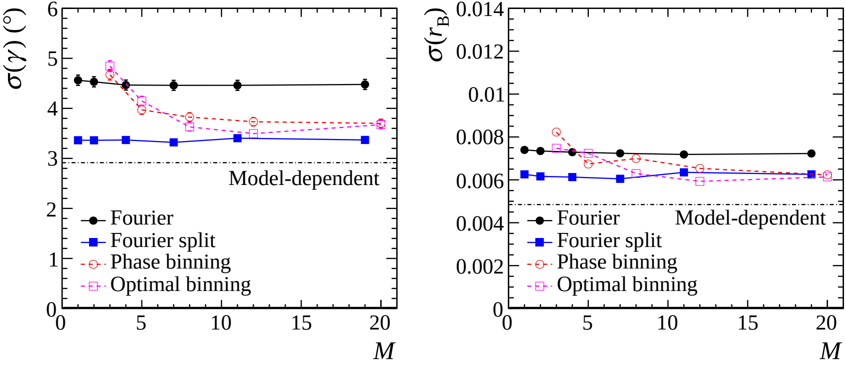
Perform Fourier analysis separately in the two regions

Can consider splitting into more regions.

More fitted parameters, but better follow $|A|$, needs optimisation.

Toy MC results: γ precision

10^7 $D \to K_S^0\pi^+\pi^-$, 10^5 $D\overline{D}$, 10^4 $B^+ \to DK^+$, 1000 toy MC samples

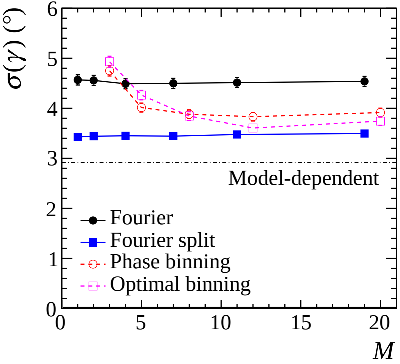


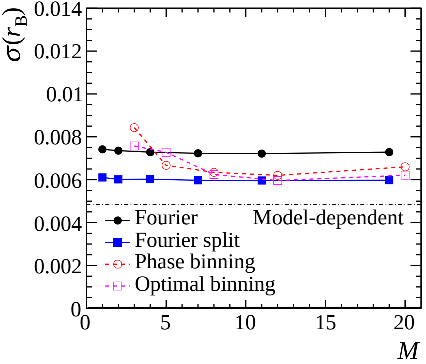
Very large sample of $D\overline{D}$ events $\Rightarrow$ precisely defined phase dependence.
Weak dependence on number of harmonics, first one dominates.*

* Not guaranteed to be the case for any D decay, or if D model used for phase differs from true one

Toy MC results: γ precision

10^7 $D \to K^0_S\pi^+\pi^-$, 10^4 $D\overline{D}$, 10^4 $B^+ \to DK^+$, 1000 toy MC samples

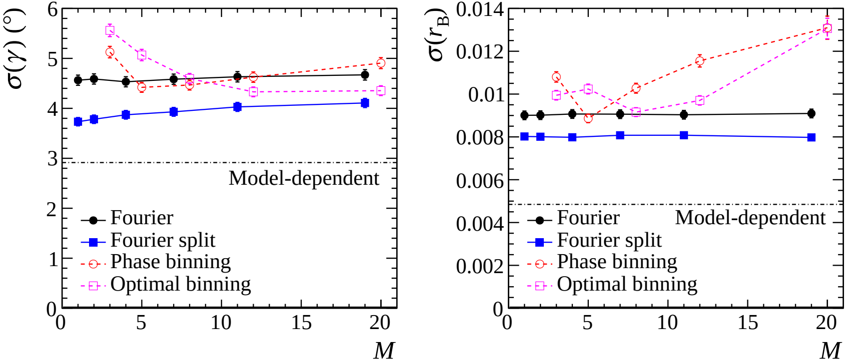




Equal samples of $D\overline{D}$ and B events ($\sim$ current situation for $B \to DK$).
$D^0\overline{D}^0$ sample does not dominate in uncertainty

Toy MC results: γ precision

10^7 $D \to K_S^0\pi^+\pi^-$, 1000 $D\overline{D}$, 10^4 $B^+ \to DK^+$, 1000 toy MC samples



Very small sample of $D\overline{D}$ events (1/10th of B sample).
Low number of free parameters of D^0 amplitude is an advantage.

Toy MC results: γ precision

Sample size	γ resolution, $^\circ$		
	Binned optimal	Fourier non-split	Fourier split
2×10^4 $B^\pm\to DK^\pm$, 10^3 $D^0\overline{D}^0$	4.33 ± 0.10	4.54 ± 0.10	3.73 ± 0.08
2×10^4 $B^\pm\to DK^\pm$, 10^4 $D^0\overline{D}^0$	3.60 ± 0.08	4.51 ± 0.10	3.43 ± 0.08
2×10^4 $B^\pm\to DK^\pm$, 10^5 $D^0\overline{D}^0$	3.49 ± 0.10	4.47 ± 0.10	3.32 ± 0.08

For reference:

- Model-dependent approach: $\sigma(\gamma)=2.91\pm0.07^\circ$
- B sample used here is roughly $\times 10$ LHCb Run I sample
- CLEO: 470 events of $D^0\overline{D}^0\to(K^0_S\pi^+\pi^-)^2$
  $(+$ other combinations, notably with $K^0_L\pi^+\pi^-)$

Possible further improvements

Fourier analysis is a good approximation to the optimum, but still not perfect: $|A_D|$ information is either ignored (non-split) or taken very roughly (split approach).

How do we choose the set of weight functions $w_n(\mathbf{z})$ such that γ precision is optimal?

Just a thought: can try to use machine learning technique:

- Choose generic $w_n(\mathbf{z})$ parametrisations with e.g. NN or BDT. (or $w_n(\phi, |A_D|, |\overline{A}_D|)$)
- Apply $w_n(\mathbf{z})$ to a set of toy MC samples, run γ fit.
- Use $|\gamma_{\mathrm{rec}} - \gamma_{\mathrm{sim}}|^2$ as cost function, and apply ML to minimise cost.

The optimal solution will likely depend on $B \to DK$ and $D^0\overline{D}^0$ samples, as well as on background levels.

Summary

- Quantum-correlated $D^0\overline{D}^0$ data are **essential for many fundamental measurements**, including CKM γ, β and charm mixing.
- These measurements have **extremely low theory uncertainties**, so any improvement in statistical treatment immediately pays off.
- Propose an approach alternative to conventional model-independent binned technique
  - Instead of performing model-independent fits in (limited number of) bins of the phase space, one works with **coefficients of Fourier series** of the phase difference spectrum.
  - Needs less parameters than binned technique: already the leading term in Fourier series contains most of the information.
    - Minimum: only 3 parameters to describe the phase variations.
  - More efficient use of QC data, could be useful when QC sample is limited, e.g. $B \to DK$ with $D \to 4h$ or $B \to DK\pi$, $D \to K^0_S\pi^+\pi^-$ double Dalitz.
- Further improvements possible, need more study.

Fourier coefficients from scattered data

Calculation of Fourier coefficients from scattered data $\phi^{(i)}$:

$$a_n = \frac{1}{\pi}\sum_{i=1}^{N}\cos(n\phi^{(i)}), \quad b_n = \frac{1}{\pi}\sum_{i=1}^{N}\sin(n\phi^{(i)}),$$

For ML fit, also need covariance matrix (uncertainties and correlations) coming from the limited sample size. It can be calculated by applying Poisson bootstrapping: each term entering the sum is multiplied by a Poisson-distributed random number with mean of 1.

E.g. dispersion is calculated from central limit theorem:

$$\sigma^2(a_n) = \frac{1}{\pi}\sum_{i=1}^{N}\cos^2(n\phi^{(i)}), \quad \sigma^2(b_n) = \frac{1}{\pi}\sum_{i=1}^{N}\sin^2(n\phi^{(i)}),$$

This is a certain approximation, but seems to work well for $N > 100$ (pulls are compatible with 1).