Statistical methods for the search for CP violation in $D^0 \to K_S^0 K^{\mp} \pi^{\pm}$ decays

Francesco Paciolla

UNIVERSITÀ DI SIENA 1240

Statistics – 17 December 2024
PhD Fisica 39° ciclo

INFN PISA

Overview

Topics:

- CP symmetry and violation
- The $D^0 \to K_S^0 K^{\mp} \pi^{\pm}$ decay
- The data sample
- The PDF
- The Amplitude Model
- Hypothesis testing (LMP, test statistic, properties)
- Point estimation (MM, estimator, properties)
- Monte Carlo

Disclaimer

This is a simplified version of the actual analysis. The following topics won't be analysed in full details

- RS and WS decay channel difference
- Systematic uncertainties
- Resonances
- Background
- Efficiency
- Parameters cross-talk

CP symmetry and its violation

CP symmetry

It is defined as the **invariance of fundamental interactions** under the combined symmetry transformations of **charge conjugation (C)** and **parity inversion (P)**

CP violation

Laws of physics are NOT the same if a **particle** is interchanged with the corresponding **antiparticle** while its spatial **coordinates are inverted**

CP symmetry and its violation

CP symmetry

It is defined as the **invariance of fundamental interactions** under the combined symmetry transformations of **charge conjugation (C)** and **parity inversion (P)**

CP violation

In short: **particles** and **antiparticles** behave differently

CP symmetry and its violation

CP symmetry

It is defined as the **invariance of fundamental interactions** under the combined symmetry transformations of **charge conjugation (C)** and **parity inversion (P)**

CP violation

In short: **particles** and **antiparticles** behave differently

Particle		Antiparticle
$D^0 \to K_S^0 K^- \pi^+$	$\neq$	$\overline{D}^0 \to K_S^0 K^+ \pi^-$

The $D^0 \to K_S^0 K^{\mp} \pi^{\pm}$ decay channel

Particle

$$D^0 \to K_S^0 K^- \pi^+$$

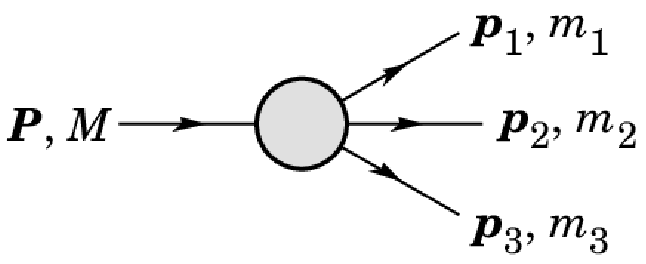
$\neq$

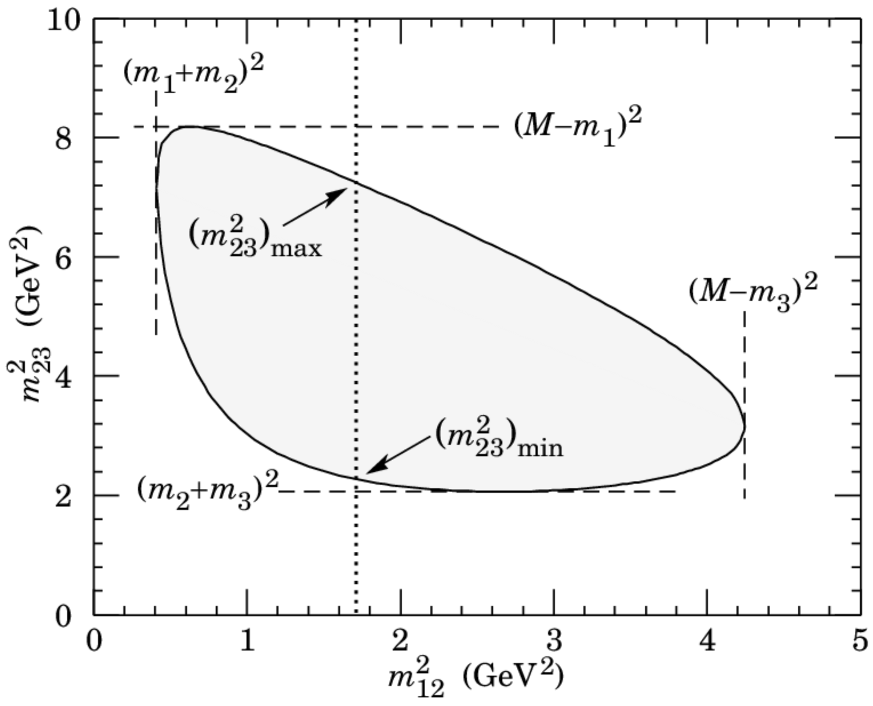
Antiparticle

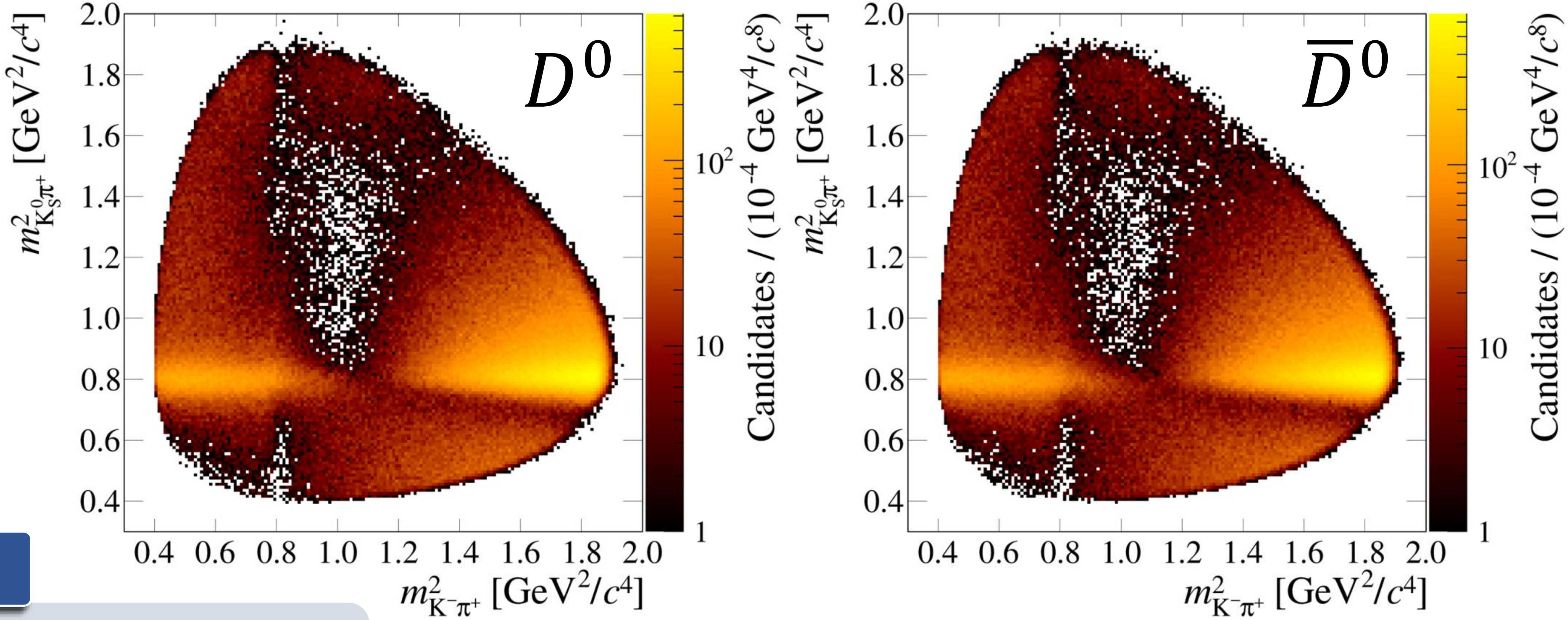
$$\bar{D}^0 \to K_S^0 K^+ \pi^-$$

Note:

- They are **three-body decays**
- Can be **displayed** by a **Dalitz Plot**
- May present **resonances**





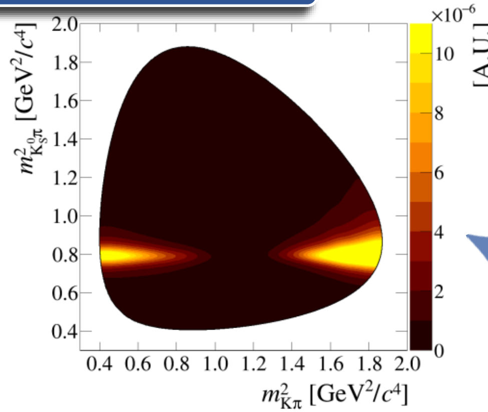


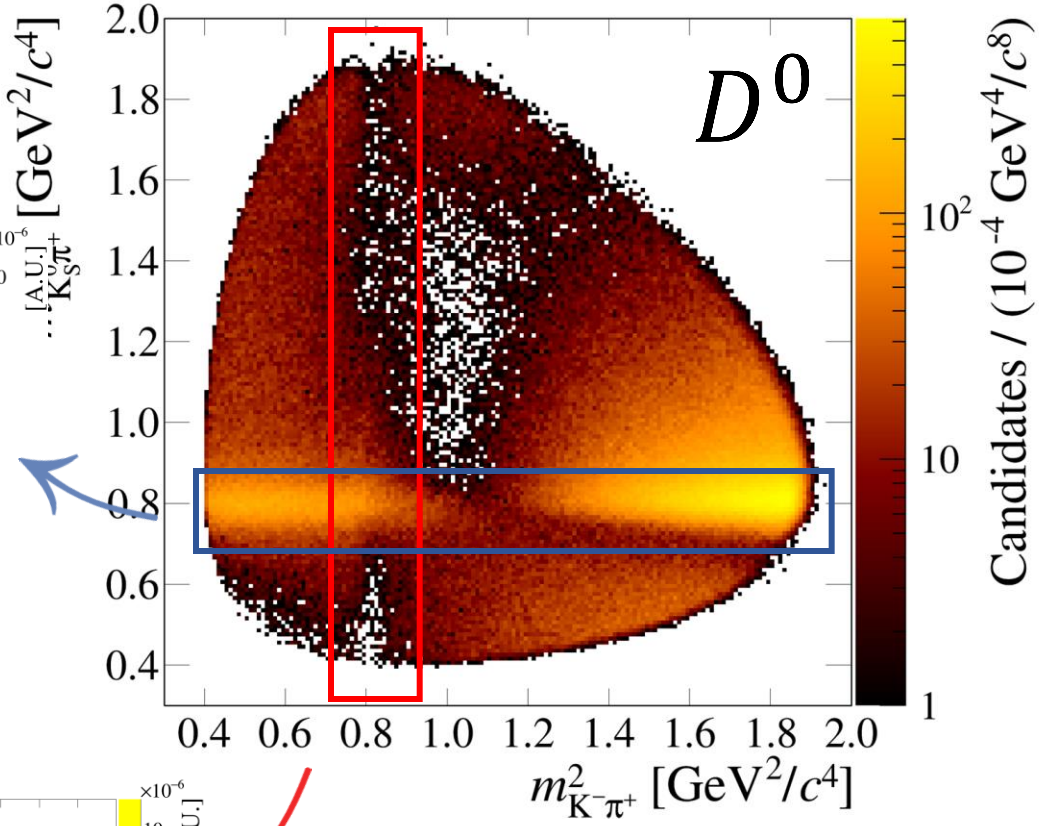
Data sample

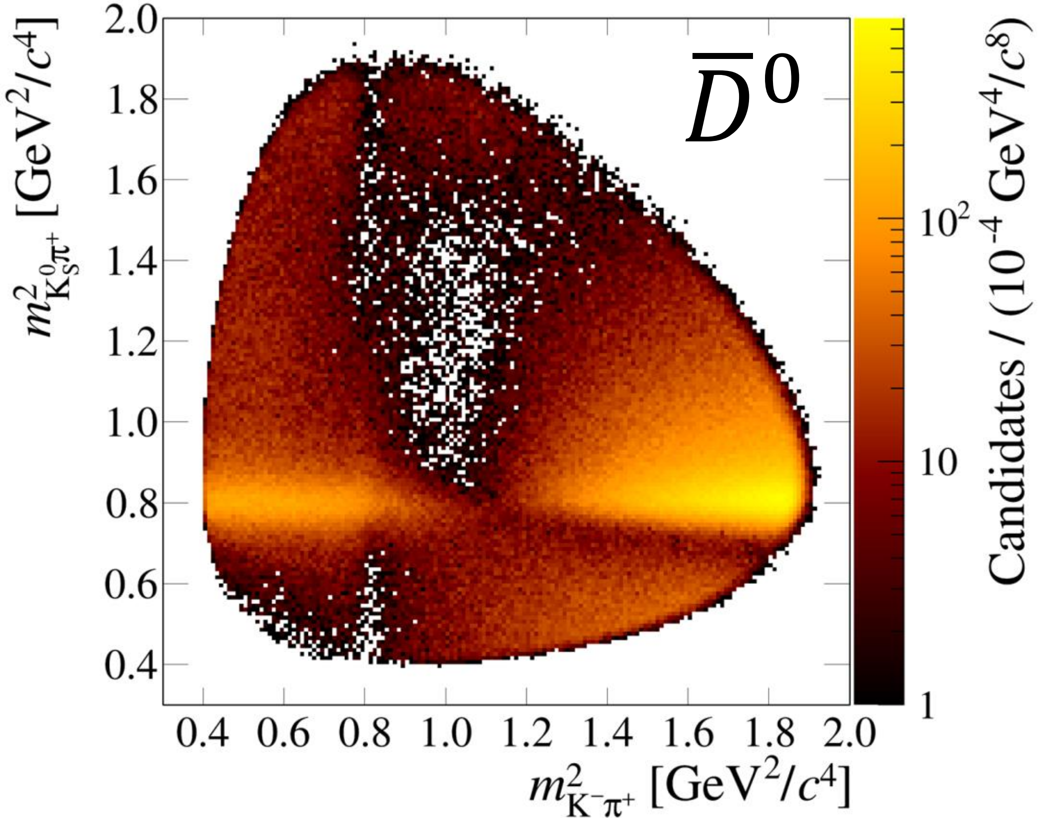
- LHCb Run 2 (2016-2018)
- $N = 900\text{K}$ events
- Background $< 4{,}5\%$
- **Toy data** (analysis is blinded)
- **Many** interfering **resonances**

The data sample

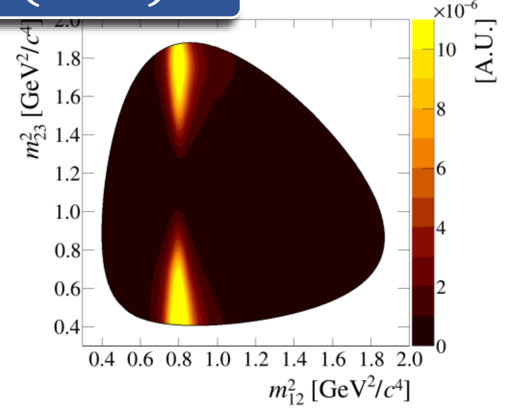
$K^*(892)^\pm$





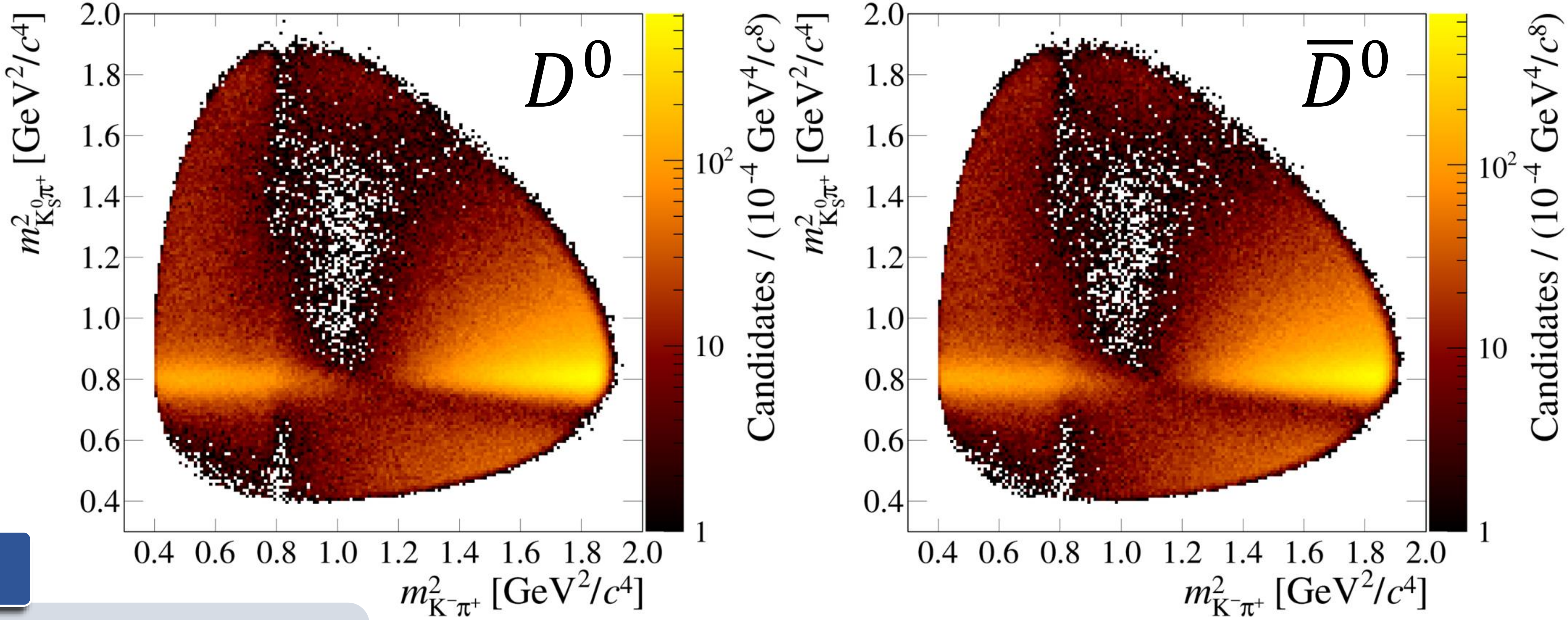


$K^*(892)^0$



$D^0 \to R(AB)\, C$		A	B	C
$D^0 \to K^0_S \overset{(-)}{K}{}^{*0}$	$\overset{(-)}{K}{}^{*0} \to K^\pm \pi^\mp$	π	K	K^0_S
$D^0 \to K^\mp K^{*\pm}$	$K^{*\pm} \to K^0_S \pi^\pm$	K^0_S	π	K
$D^0 \to \pi^\mp(\rho^\pm, a^\pm)$	$(\rho^\pm, a^\pm) \to K^0_S K^\pm$	K	K^0_S	π

The data sample



Data sample

- LHCb Run 2 (2016-2018)
- $N = 900\text{K}$ events
- Background $< 4{,}5\%$
- **Toy data** (analysis is blinded)
- **Many** interfering **resonances**

Goal

Search for any **statistically significant differences** between the two data sets that are **related to CP violation**

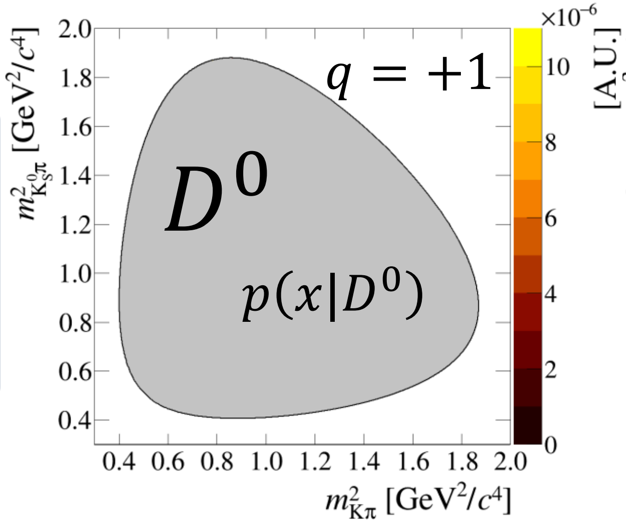
PDF needed

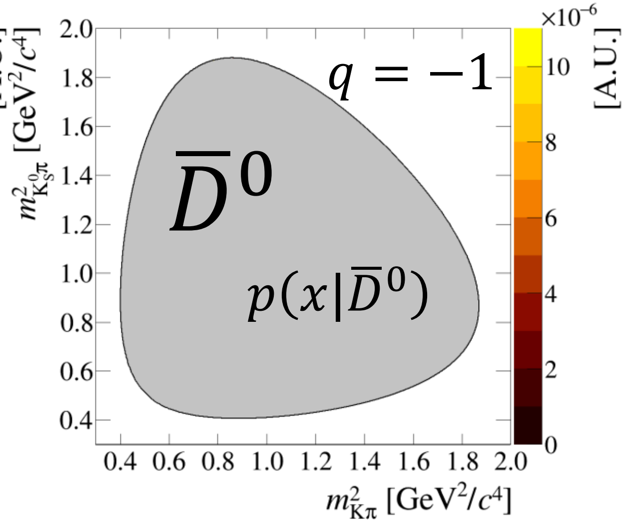
The PDF: derivation

PDF

$$p_{\{\pm1\},\Omega}(X) = \ldots ?$$

$$X = (q, x) \quad \begin{cases} q = \text{ flavour } D^0, \bar{D}^0 & q \in \{\pm1\} \\ x = \text{ Dalitz plot position} & x \in \Omega \end{cases}$$

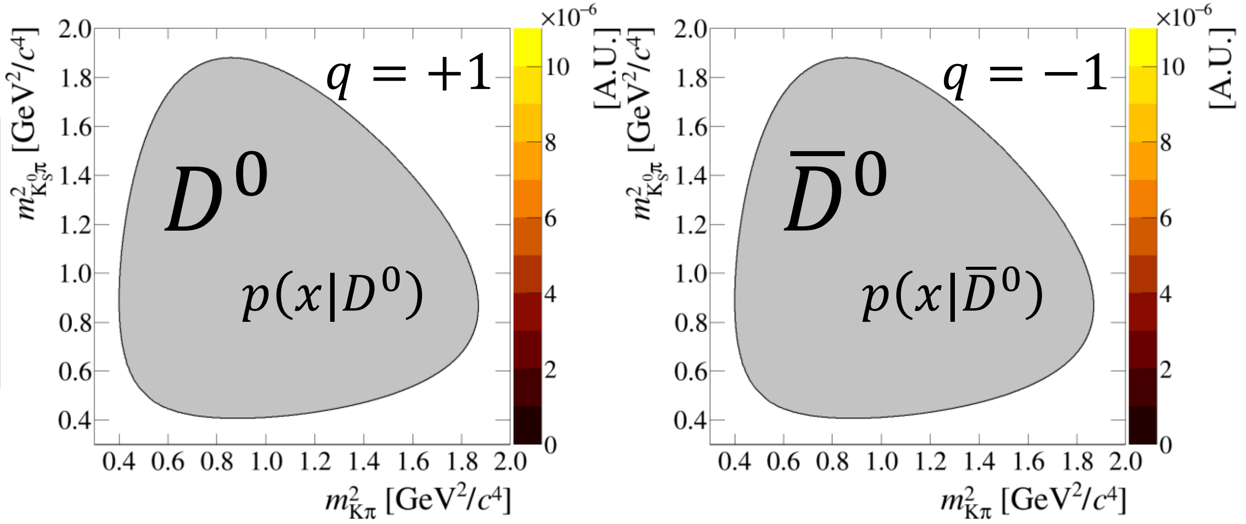




PDF

$$p_{\{\pm1\},\Omega}(X) = \{\alpha p(x|D^0)\}^{\frac{1+q}{2}}\{(1-\alpha)p(x|\bar{D}^0)\}^{\frac{1-q}{2}}$$

$$X = (q,x) \quad \begin{cases} q = \text{ flavour } D^0, \bar{D}^0 & q \in \{\pm1\} \\ x = \text{ Dalitz plot position} & x \in \Omega \end{cases}$$



Proof

$$\begin{aligned} p_{\{\pm1\},\Omega}(X) &= p_{\{\pm1\},\Omega}(q,x) = p(q \in \{\pm1\} \cap x \in \Omega) = \\ &= P_{\{\pm1\}}(q)p_\Omega(x|q) = \\ &= \begin{cases} P(D^0)p(x|D^0) & q = +1 \\ P(\bar{D}^0)p(x|\bar{D}^0) & q = -1 \end{cases} = \\ &= \begin{cases} \alpha p(x|D^0) & q = +1 \\ (1-\alpha)p(x|\bar{D}^0) & q = -1 \end{cases} = \\ &= \boldsymbol{\{\alpha p(x|D^0)\}^{\frac{1+q}{2}}\{(1-\alpha)p(x|\bar{D}^0)\}^{\frac{1-q}{2}}} \end{aligned}$$

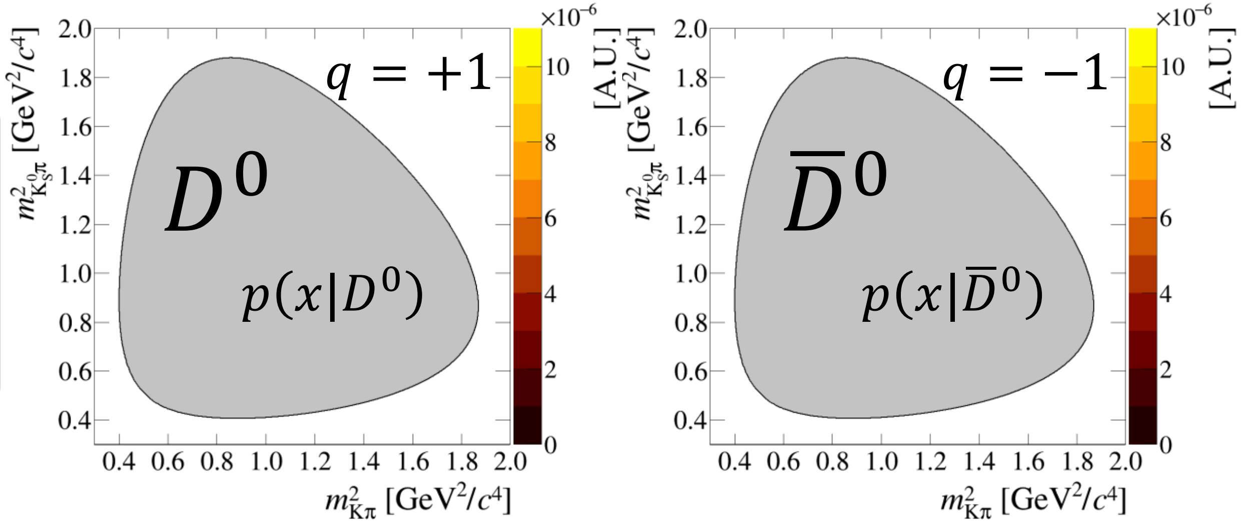
- $P(A,B) = P(A \cap B)$ joint probability
- $P(B|A) = \frac{P(A,B)}{P(A)}$ conditional probability
- $q = +1 \to D^0$, $q = -1 \to \bar{D}^0$
- $P(D^0) = \alpha, \quad P(\bar{D}^0) = 1-\alpha$
  α = global asymmetry $\quad \alpha \in [0,1]$
- Rearrange

The PDF

PDF

$$p_{\{\pm1\},\Omega}(X) = \{\alpha p(x|D^0)\}^{\frac{1+q}{2}}\{(1-\alpha)p(x|\bar{D}^0)\}^{\frac{1-q}{2}}$$

$$X = (q,x) \quad \begin{cases} q = \text{ flavour } D^0, \bar{D}^0 & q \in \{\pm1\} \\ x = \text{ Dalitz plot position} & x \in \Omega \end{cases}$$



α

The red terms can be interpreted as a **Bernoulli distribution** where α is the success probability

$$f(q;\alpha) = \alpha^{\frac{1+q}{2}}(1-\alpha)^{\frac{1-q}{2}}$$

$$P(D^0) = \alpha \quad P(\bar{D}^0) = 1-\alpha$$

$p(x|q)$

The green terms are the **flavour specific PDF** of the data Dalitz plot distribution

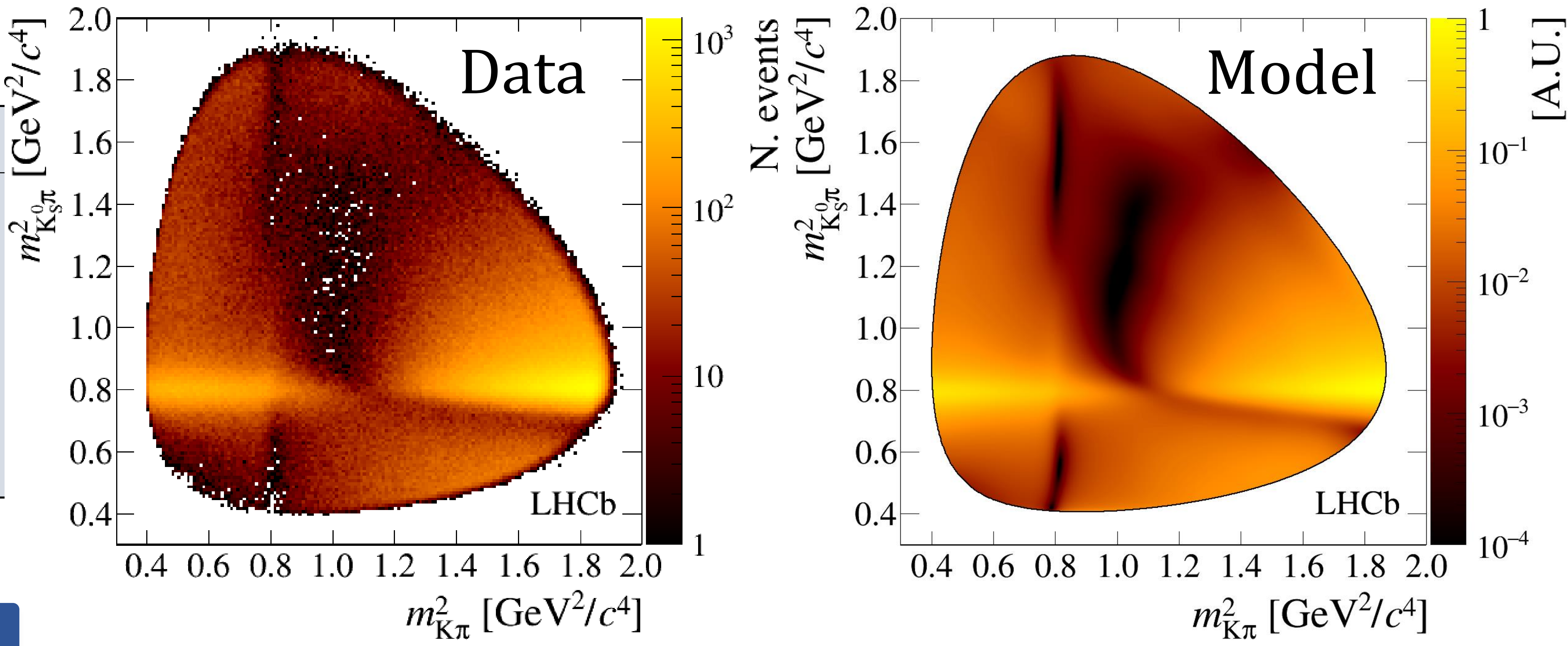
$$p_\Omega(x|q=+1) = p(x|D^0)$$

$$p_\Omega(x|q=-1) = p(x|\bar{D}^0)$$

To be defined

The PDF: amplitude model from LHCb Run 1

$R \to K\pi$	$R \to K_S^0 K^\pm$
$K^*(892)^0$	$a_0(892)^-$
$K^*(892)^\pm$	$a_2(1320)^+$
$K_1^*(1410)^{0,\pm}$	$a_0(1450)^\pm$
$K_0^*(1430)^{0,\pm}$	$\rho(1450)^\pm$
$K_2^*(1430)^0$	$\rho(1700)^\pm$



Isobar model

$$A_{K_S^0 K^\mp \pi^\pm}(x) = \varepsilon(x) \left| \sum_R a_R e^{i\phi_R} \mathcal{M}_R(x) \right|^2$$

- $x = (m^2_{K\pi}, m^2_{K_S^0\pi})$ position in the Dalitz plot
- The sum runs over 2-body intermediate resonances R
- $\mathcal{M}_R(x)$ is the matrix element, $\varepsilon(x)$ is the efficiency model

LHCb collaboration, R. Aaij et al., "*Studies of the resonance structure in $D^0 \to K_S^0 K^\mp \pi^\pm$ decays*", [PRD.93.052018]

The PDF: flavour specific $PDFs$

CPV parameters

The amplitude model can be re-parametrized using

$$p_{\Omega}(x|q) \equiv p_{\pm}(x;\vec{\theta}) = \varepsilon(x)\left|\sum_{R} a_R(1 \pm \boldsymbol{\Delta a_R})e^{i(\phi_R \pm \boldsymbol{\Delta\phi_R})}\mathcal{M}_R(x)\right|^2 \qquad \begin{cases} D^0 \to + \\ \bar{D}^0 \to - \end{cases}$$

So, $\vec{\boldsymbol{\theta}} = (\boldsymbol{\Delta a}, \boldsymbol{\Delta\phi})_{\boldsymbol{R}}$ are the set of ***CP*-violating parameters**.

- A search for CPV must deal with these $2N_R$ parameters
- A_{CP} is a function of $\vec{\theta} = (\Delta a, \Delta\phi)_R$

Linear approximation

Since ***CPV* is expected to be small**, the $\boldsymbol{p_{\pm}(x;\vec{\theta})}$ **can be expanded in** $\vec{\boldsymbol{\theta}} = \vec{\boldsymbol{0}}$ as

$$p_{\pm}(x;\vec{\theta}) = f(x) \pm \sum_{\chi\in\vec{\theta}} \chi g_{\chi}(x) + O(\vec{\theta}^2)$$

where

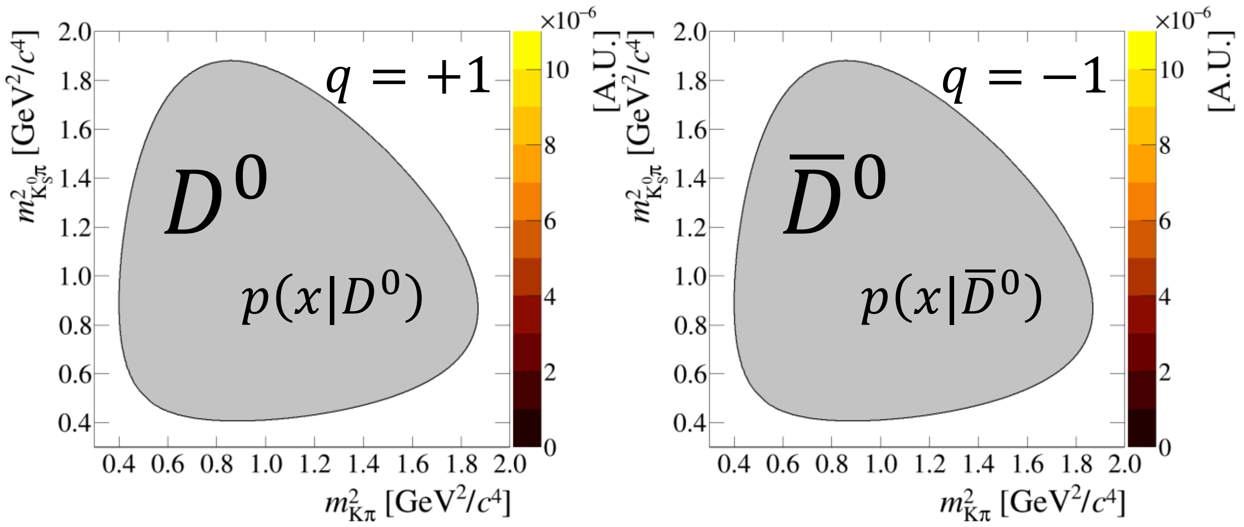
$$f(x) \equiv p_{\pm}(x;\vec{\theta})\Big|_{\vec{\theta}=\vec{0}} = p(x;\vec{0}) \qquad g_{\chi}(x) \equiv \frac{\partial}{\partial\chi}p_{\pm}(x,\vec{\theta})\Big|_{\vec{\theta}=\vec{0}}$$

The PDF: final form

PDF

$$p(q,x;\alpha,\vec{\theta}) = \{\alpha p_+(x;\vec{\theta})\}^{\frac{1+q}{2}} \{(1-\alpha)p_-(x;\vec{\theta})\}^{\frac{1-q}{2}}$$

- q = flavour $D^0, \overline{D}^0$ and $q \in \{\pm 1\}$
- x = Dalitz plot position and $x \in \Omega$
- α = global asymmetry and $\alpha \in [0,1]$
- $p_\pm(x;\vec{\theta}) = f(x) \pm \sum_{\chi \in \vec{\theta}} \chi g_\chi(x)$ linear approximation



Goal

The **search for CP violation** may be carried out through:

- **Hypothesis testing**
- **Point estimation** of $\vec{\theta} = (\Delta a, \Delta\phi)_R$

Hypothesis test

Hypothesis testing: LMP test

Hp. test

The search for CPV can interpreted as a **hypothesis test**

- $H_0 = CP$ conservation $\quad \vec{\theta} = \vec{0}$
- $H_1 = CP$ violation $\quad \vec{\theta} \neq \vec{0}$

Considering that:

- $\boldsymbol{H_1}$ **depends on** $\boldsymbol{\vec{\theta}}$ and is not simple
- Predictions say $\boldsymbol{CPV \sim 0}$

the most convenient test is the **Locally Most Powerful** (**LMP**) test

LMP test

- It can be used when $\boldsymbol{H_0 \sim H_1}$
- It is the **most powerful test** $\forall \vec{\theta} \in I(\vec{\theta}_0)$
- For a data set $\vec{X}$ of N observations, **the test statistic** is

$$t_\chi(\vec{X}) = s(\chi)\Big|_{\vec{\theta}=\vec{0}} = \frac{\partial}{\partial \chi} \log \mathcal{L}_{\vec{X}}(\vec{\theta})\Big|_{\vec{\theta}=\vec{0}} \qquad \forall \chi \in \vec{\theta}$$

Hypothesis testing: test statistic

Likelihood

$$\mathcal{L}_X(\vec{\theta},\alpha) = \{\alpha\mathcal{L}_+(\vec{\theta}|x)\}^{\frac{1+q}{2}}\{(1-\alpha)\mathcal{L}_-(\vec{\theta}|x)\}^{\frac{1-q}{2}}$$

- $\mathcal{L}_\pm(\vec{\theta}|x) = f(x) \pm \sum_{\chi\in\vec{\theta}} \chi g_\chi(x)$ linear approximation
- $\vec{\theta} \in I(\vec{\theta}_0 = \vec{0})$ and $\alpha \in [0,1]$

$t_\chi(\vec{X})$

$$t_\chi(\vec{X}) = s(\chi)\Big|_{\vec{\theta}=\vec{0}} = \frac{\partial}{\partial\chi}\log\mathcal{L}_{\vec{X}}(\vec{\theta})\Big|_{\vec{\theta}=\vec{0}} =$$

$$= \frac{\partial}{\partial\chi}\ln\prod_{i=1}^{N}\left[\alpha\mathcal{L}_+(\vec{\theta}|x_i)\right]^{\frac{1+q_i}{2}}\left[(1-\alpha)\mathcal{L}_-(\vec{\theta}|x_i)\right]^{\frac{1-q_i}{2}}\Bigg|_{\vec{\theta}=\vec{0}} =$$

$$= \sum_{i=1}^{N}\left(\frac{1+q_i}{2}\right)\frac{g_\chi(x_i)}{f(x_i)+\sum_{\psi\in\vec{\theta}}\psi g_\psi(x_i)} - \left(\frac{1-q_i}{2}\right)\frac{g_\chi(x_i)}{f(x_i)-\sum_{\psi\in\vec{\theta}}\psi g_\psi(x_i)}\Bigg|_{\vec{\theta}=\vec{0}} =$$

$$= \boldsymbol{\sum_{i=1}^{N} q_i \frac{g_\chi(x_i)}{f(x_i)} \qquad \forall\chi\in\vec{\theta}}$$

Hypothesis testing: properties

$t_\chi(\vec{X})$

The test statistic $t_\chi(\vec{X})$ is

$$t_\chi(\vec{X}) = \sum_{i=1}^{N} q_i \frac{g_\chi(x_i)}{f(x_i)}$$

$t_{thr}(\alpha)$

At fixed α, the threshold $t_{thr}(\alpha)$ is given by

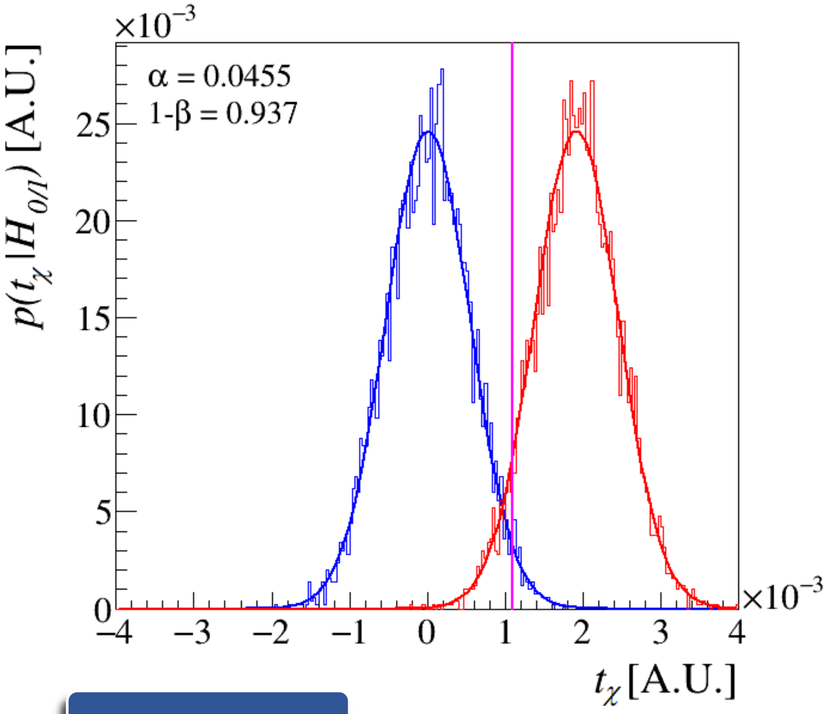
$$\int_{t_{thr}(\alpha)}^{+\infty} p(t_\chi|H_0)\, dt_\chi = \alpha$$

α is the significance level

$\text{pow}(\chi)$

At given χ, he power $\text{pow}(\chi)$ is given by

$$\text{pow}(\chi) = \int_{t_{thr}(\alpha)}^{+\infty} p(t_\chi|H_1)\, dt_\chi$$



Example

- $\chi = \Delta\phi_{K*(892)^0}$
- $p(t_\chi|H_0) \to \chi_0 = 0°$
- $p(t_\chi|H_1) \to \chi_1 = 1.5°$

Point estimation

Point estimation: MM estimator

MLE vs MM

Given the complexity of the $p(q, x; \alpha, \vec{\theta})$

- the Maximum Likelihood Estimator (**MLE**) is **unfeasible**
- the Method of Moments (**MM**) could be **useful**

Method of Moments

In its most general form, **the MM states that**

$$\text{if} \quad \forall \chi \in \vec{\theta} \quad \begin{cases} \exists \mathbf{Var}[\boldsymbol{t}_\chi(\boldsymbol{x})] \text{ and it is finite} \\ h_\chi(\vec{\theta}) \equiv \mathbf{E}[\boldsymbol{t}_\chi(\boldsymbol{x})] \\ \exists h_\chi^{-1} \text{ inverse of } h_\chi(\vec{\theta}) \text{ at least } \forall \vec{\theta} \in I(\vec{\theta}_0) \end{cases}$$

then any $\chi \in \vec{\theta}$ can be estimated using the **estimator** defined as

$$\hat{t}_\chi(\vec{X}) = h^{-1}\left(\frac{1}{N}\sum_{i=1}^{N} t_\chi(X_i)\right)$$

Point estimation: MM estimator

Method of Moments

In its most general form, **the MM states that**

$$\text{if} \quad \forall \chi \in \vec{\theta} \quad \begin{cases} \exists \mathbf{Var}\left[\boldsymbol{t}_\chi(\boldsymbol{x})\right] \text{ and it is finite} \\ h_\chi(\vec{\theta}) \equiv \mathbf{E}\left[\boldsymbol{t}_\chi(\boldsymbol{x})\right] \\ \exists h_\chi^{-1} \text{ inverse of } h_\chi(\vec{\theta}) \text{ at least } \forall \vec{\theta} \in I(\vec{\theta}_0) \end{cases}$$

then any $\chi \in \vec{\theta}$ can be estimated using the **estimator** defined as

$$\hat{t}_\chi(\vec{X}) = h^{-1}\left(\frac{1}{N}\sum_{i=1}^{N} t_\chi(X_i)\right)$$

MLE vs MM

Given the complexity of the $p(q, x; \alpha, \vec{\theta})$

- the Maximum Likelihood Estimator (**MLE**) is **unfeasible**
- the Method of Moments (**MM**) could be **useful**

LMP statistic

Let's look at the **properties** of the **LMP statistic**

- $t_\chi(X) = q\dfrac{g_\chi(x)}{f(x)}$
- $\mathrm{E}[t_\chi] = A_{\chi\psi}\theta_\psi$
- $\mathrm{Cov}[t_\chi, t_\psi] = A_{\chi\psi} + (2\alpha - 1)B_{\chi\psi\xi} - \mathrm{E}[t_\chi]\mathrm{E}[t_\psi]$

MM is satisfied

The coefficients are:

$$A_{\chi\psi} \equiv \int_\Omega \frac{g_\chi(x) g_\psi(x)}{f(x)} dx$$

$$B_{\chi\psi\xi} \equiv \int_\Omega \frac{g_\chi(x) g_\psi(x)}{f(x)} dx$$

Point estimation: estimator

Estimator

Defined $\mathbb{A} \equiv \left(A_{\chi\psi}\right)_{\chi,\psi\in\vec{\theta}}$ and $\mathbb{B} \equiv \left(B_{\chi\psi\xi}\right)_{\chi,\psi,\xi\in\vec{\theta}}$, the **MM estimator** is given by

$$\boldsymbol{\hat{t}_\chi(\vec{X}) = \frac{1}{N}\mathbb{A}^{-1}_{\chi\psi}\sum_{i=1}^{N} q_i \frac{g_\psi(x_i)}{f(x_i)}}$$

- $\mathrm{E}\left[\hat{t}_\chi\right] = \theta_\chi$
- $\mathrm{Cov}\left[\hat{t}_\chi, \hat{t}_\psi\right] = \frac{1}{N}\mathbb{A}^{-1}_{\chi\psi}\left[\mathbb{A}_{\chi\psi} + (2\alpha - 1)\mathbb{B}_{\chi\psi\xi}\theta_\xi\right]\left(\mathbb{A}^{-1}_{\chi\psi}\right)^T - \frac{1}{N}\theta_\chi\theta_\psi$

Properties

The estimator is

- Asymptotically normally distributed
- Consistent
- Unbiased
- Highly efficient $(\varepsilon > 91\%)$
- Properties proved analytically

Good Estimator!

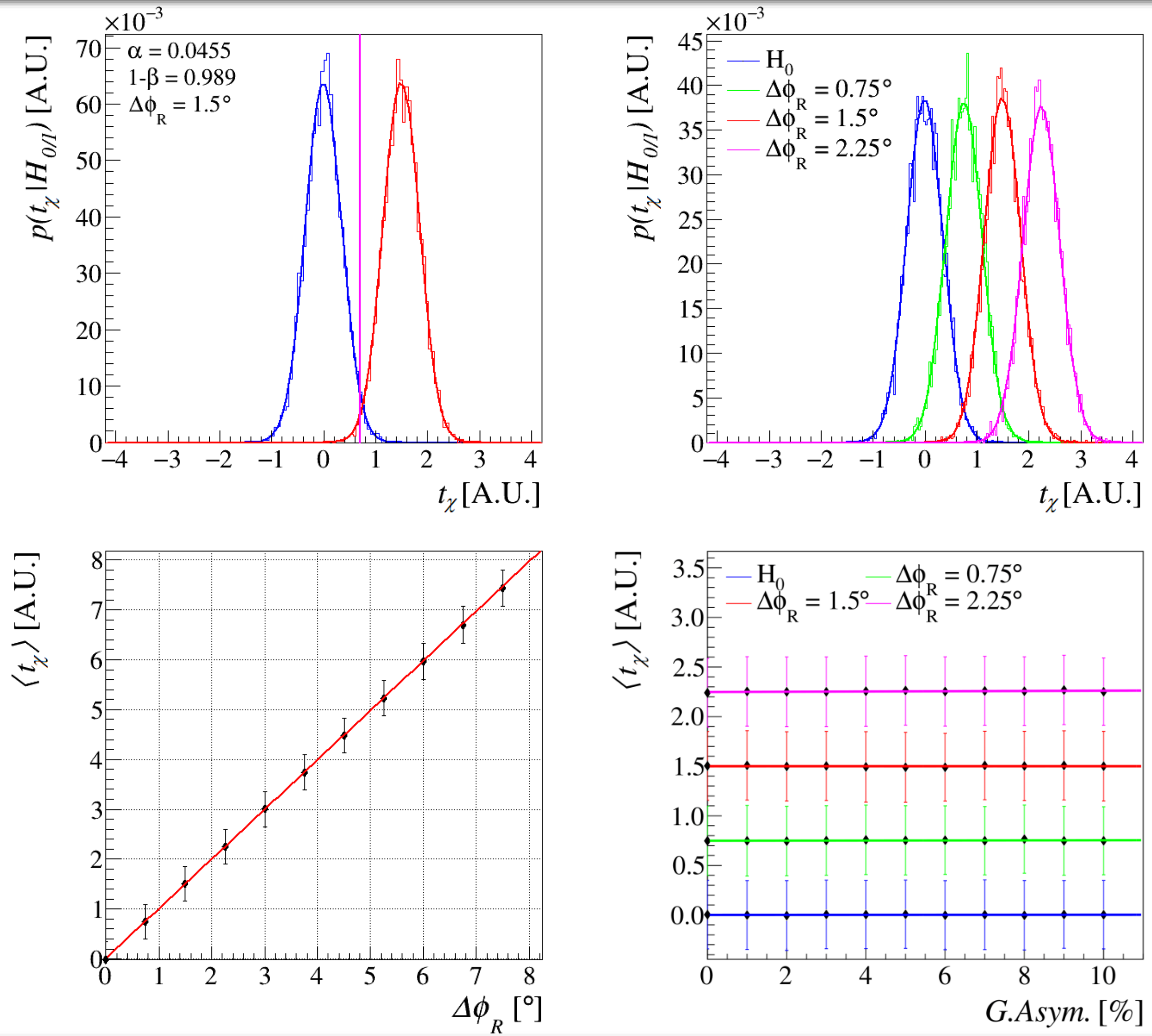
Point estimation: Monte Carlo simulation

Method

1. Inject CPV in the model through a selected $\chi \in \vec{\theta}$
2. Generate two random samples $S_+(D^0)$ and $S_-(\overline{D}^0)$
3. Evaluate $\hat{t}_\theta$ using its definition
$$\hat{t}_\chi(\vec{X}) = \frac{1}{N}\mathbb{A}^{-1}_{\chi\psi}\sum_{i=1}^{N} q_i \frac{g_\psi(x_i)}{f(x_i)}$$
4. Repeat 5000 times to obtain $p\left(\hat{t}_\chi; \alpha, \vec{\theta}\right)$
5. Vary χ and α over the expected theoretical and experimental ranges

Example

- $\chi = \Delta\phi_{K*(892)^0}$
- Same results for other parameters



Point estimation: Monte Carlo Simulation

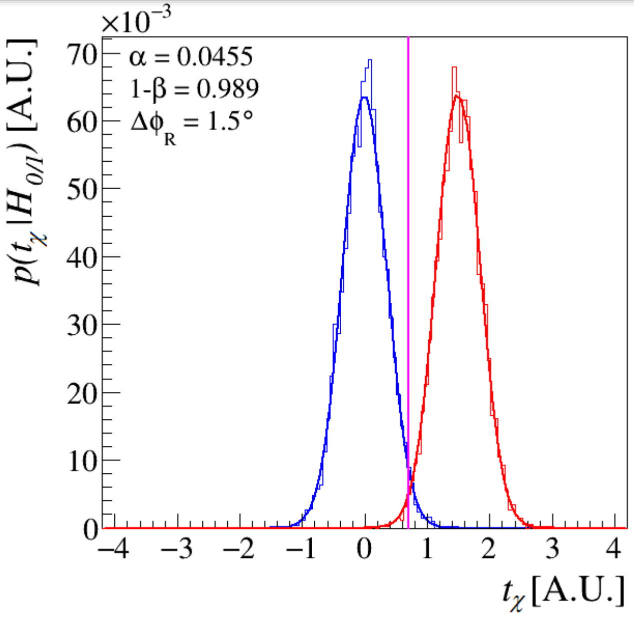
Results

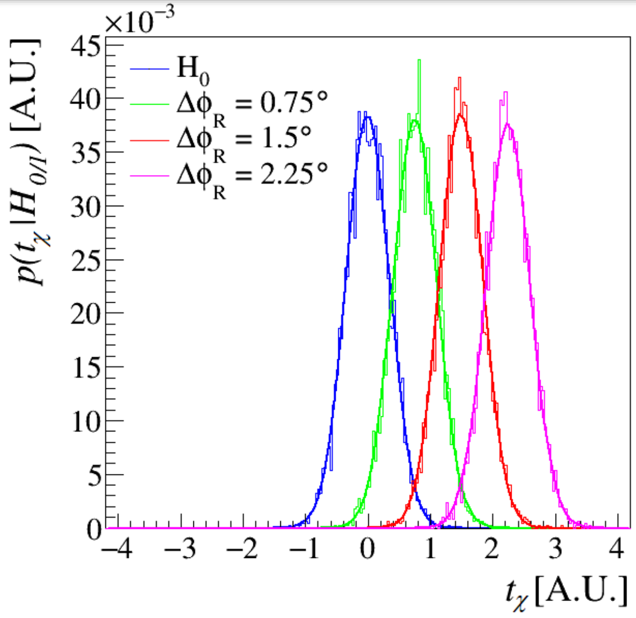
The numerical analysis validates the estimator's properties:

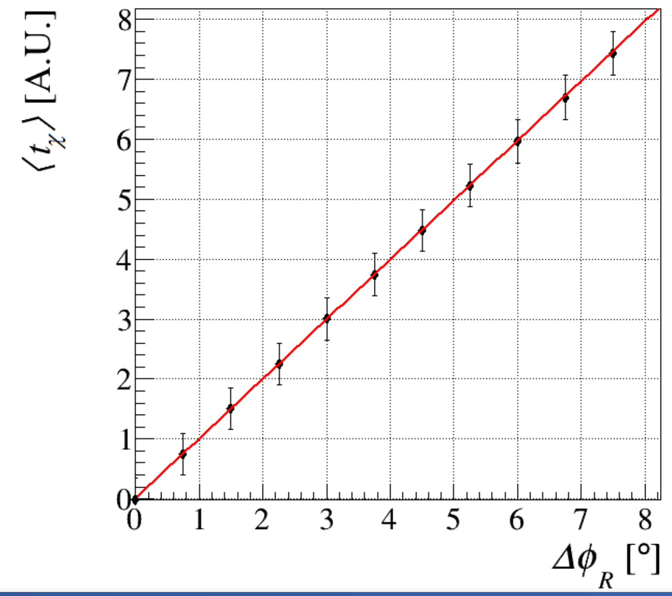
- The **asymptotic normality** of $p(\hat{t}_\chi; \alpha, \vec{\theta})$ implies $\hat{t}_\chi$ **consistency**
- The **linear** fit coincides with the bisector. This implies $\hat{t}_\chi$ **unbiasedness**
- The variance obtained from the fit is equal to the analytical prediction. The $\hat{t}_\chi$ **high efficiency** is confirmed
- The $\mathrm{E}[\hat{t}_\chi]$ shows **no dependence from the global asymmetry** $\boldsymbol{\alpha}$

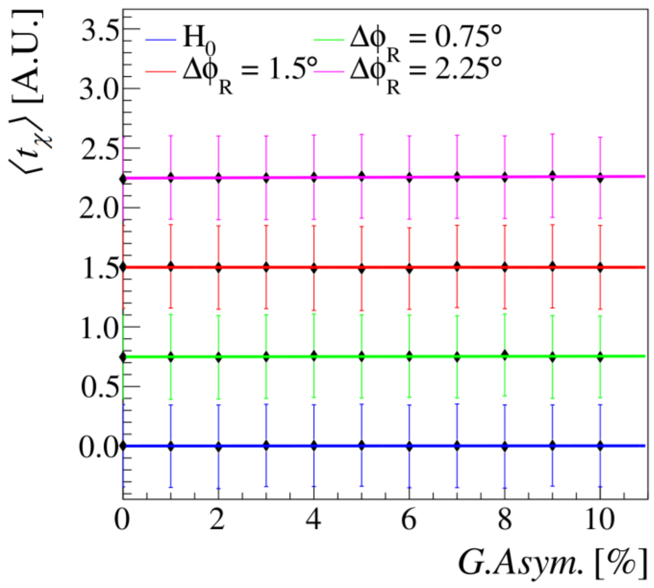
Example

- $\chi = \Delta\phi_{K*(892)^0}$
- Same results for other parameters









Thank you for your attention