Results for approximate profiled FC for $B^0 \to K^\star \mu\mu$

or "How we spent our xmas vacation"

Alessio Boletti, **Stefano Lacaprara**

stefano.lacaprara@pd.infn.it

INFN Padova

no meeting,
work, January 9, 2017

Introduction

- After discussion with StatCom, the outcome was that **none** of the many methods we proposed to estimate the statistical uncertaities are good enough;
  - the proposed solution is to build the Neyman construction
  - due to the slowness of our fit, the full construction is simply impossible
  - instead, we agreed upon a 1D FC approach by profiling the likelihood on data at fixed P_1 (P'_5) values.
  - here will describe the procedure and the results;
- CAVEAT! most of the work has been done by Alessio, which is now in well deserved vacation. He described the full procedure to me and I started from his work, but I might have missed something...

procedure description in the PAS

To ensure correct coverage for the uncertainties of the angular observables, the Feldman-Cousins method [?] is used with nuisance parameters. Two main sets of pseudo-experimental samples are generated to compute the coverage for the two angular obervables P_1 and P'_5, respectively. The first (second) set, used to compute the coverage for P_1 (P'_5), is generated by assigning values to the other observables as obtained by profiling the likelihood on data at fixed P_1 (P'_5) values. When fitting the pseudo-experimental samples the same fit procedure as in data is applied.

Procedure description

- We start from the 2D likelihood computed on data as a function of P_1 and P_5'
- taking into account the physical boundaries
  - The $\mathcal{L}$ is computed in each point of the grid fixing P_1 and P_5' and maximizing the $\mathcal{L}$ for yields and A_5^s
  - Or the minimization failed or it is outside the pyhisical region (which depends on A_5^s)
- Then we profile it vs P_1 and P_5', respectively
  - *e.g.* suppose we want to check on P_1: then P_5' is the one which minimize the profiled likelihood.
  - We start from a reasonable point for P_1, close to the $\Delta \log \mathcal{L} = 0.5$ line for the data.
  - if we hit a physical boundary, the minimum can be along the boundary itself
- Then we generate toys using as input parameters P_1 and P_5'.
- We generate 100 toys (data-like size)

Procedure description (cont'ed)

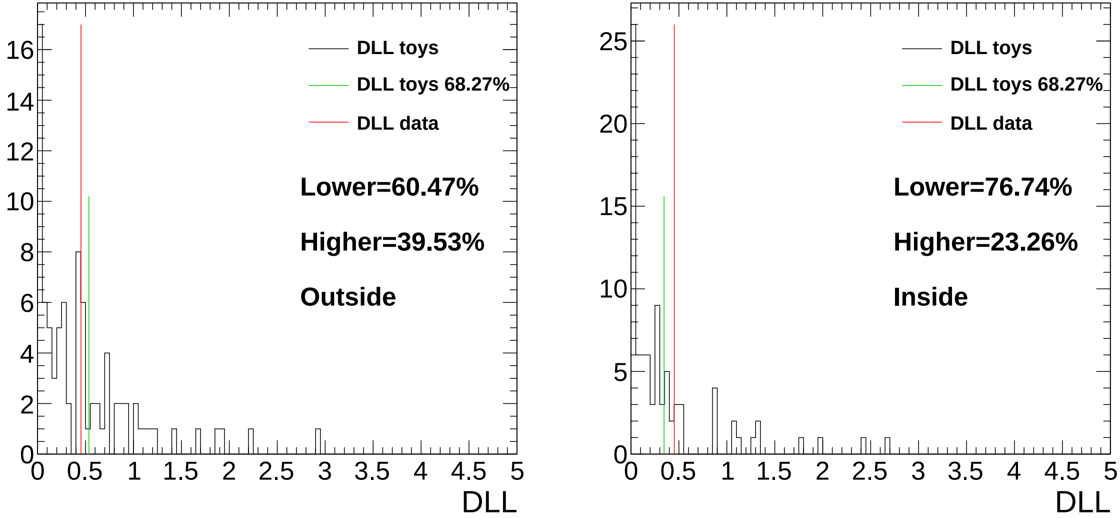
Each toy is fitted with the full pdf as done for data

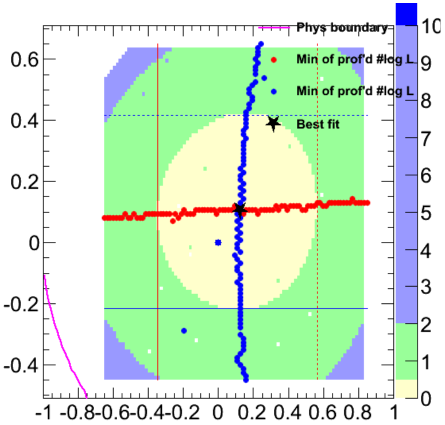
- we repeat the fit with 20 different set of 20 initial values of P_1 and P_5'
- the 20 points are choosen randomly a 2D gaussian distribution around a central value as follow
  - we try a fully free fit (leaving also P_1 and P_5')
  - if the fit converge, then we use the best fit value;
  - if not the central point is the generated one.
- the results of the 20 fits provide 20 likelihood values in 20 sets of (P_1, P_5'),
- to find the absolute minimum, we fit the 20 values with a 2D gaus function
  - Need some more detail on the 2D gauss fit
- Eventually, we have 100 toys, and 100 values for the likelihood.

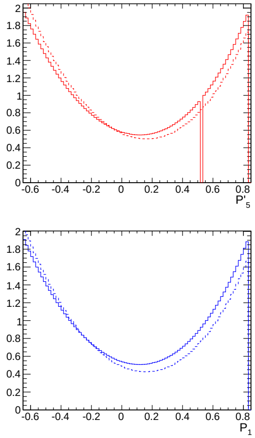
Procedure description (cont'ed)

- We compute for each of the 100 toys the $\Delta \log \mathcal{L}$ between the toy best fit (min of the 2D gaus fit) and the max value of the $\mathcal{L}$ computed along the profile
  - that is with the value of $\mathcal{L}$ of the gerenated point.
- We compute $\Delta \log \mathcal{L}$ of the data as the difference between the $\mathcal{L}$ computed in the data best fit with that computed in the point used to generate the toy.
- We compute how many times the DLL(toy) is greater that DLL(data), and divided by the total of toys
  - In principle there should be 100 toys, this is not true for two reason:
    - ★ the 2D gaus fit fails: to be looked at.
    - ★ problem with the job (batch-system related): now are very few
    - ★ The yellow hist in the following page show the number of sucessfull toys
- If the ratio is $> 68.27\%$ then the toy generation point is outside the 1σ boundary for data, otherwise it's inside.
- the ratio is shown in the 4 graphs in next slides
- We repeat the procedure to compute the lower/higher uncertainties for both P_1 and P_5' (4 "directions")

Example


DLL toys
DLL toys 68.27%
DLL data
Lower=60.47%
Higher=39.53%
Outside
DLL
DLL toys
DLL toys 68.27%
DLL data
Lower=76.74%
Higher=23.26%
Inside
DLL



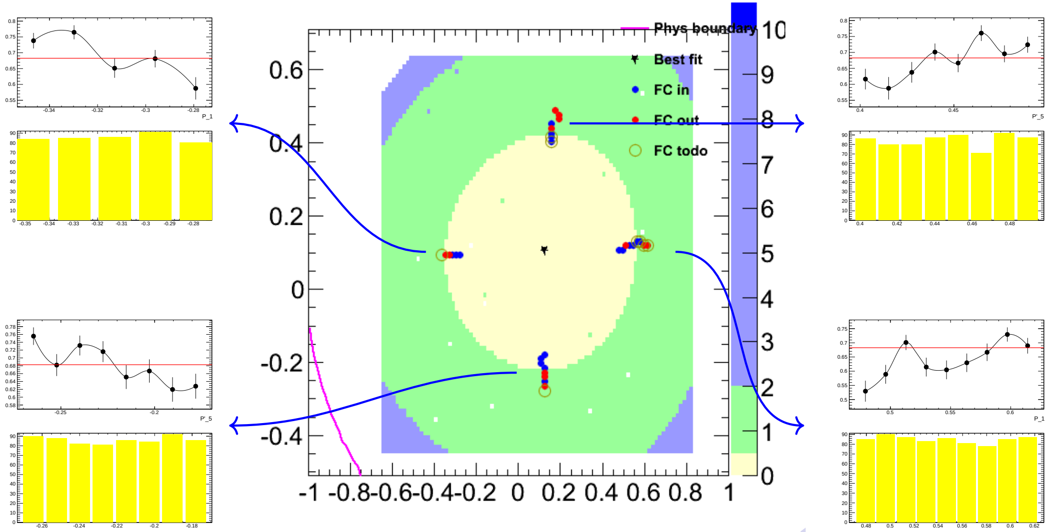


LEFT:

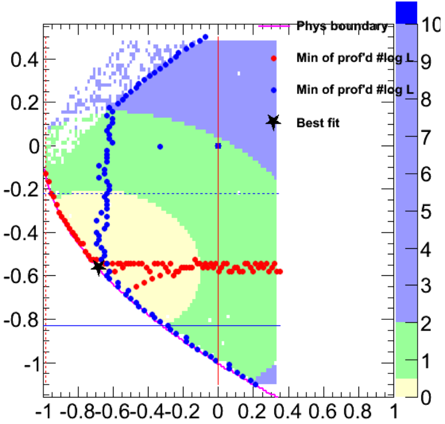
- 2D $\log\mathcal{L}$ for data
- $\star$ best fit
- magenta: physical limit
- red dots: min profiled $\log\mathcal{L}$ along P_1
- blue dots: likewise for P_5'

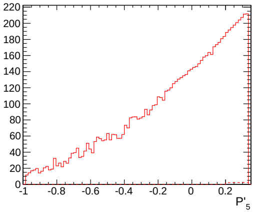
RIGHT

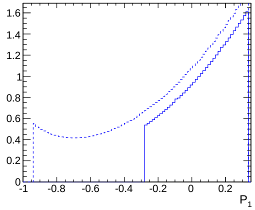
- profiled $\log\mathcal{L}$ for roughly $\log\mathcal{L}\sim 0.5$
- P_1 (red solid/dash)
- P_5' (blue solid/dash)


Phys boundary
Best fit
FC in
FC out
FC todo
P_1
P'_5

Bin 1







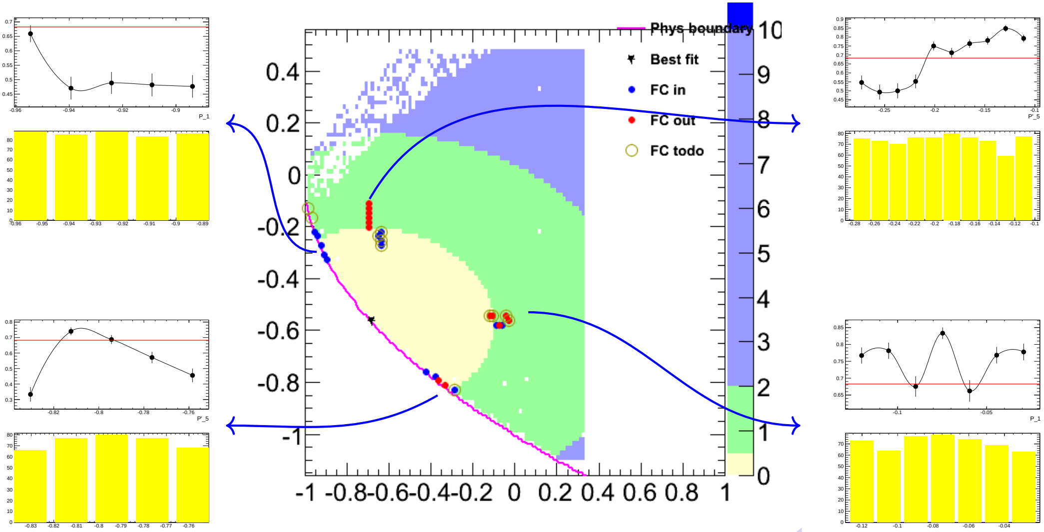
LEFT:

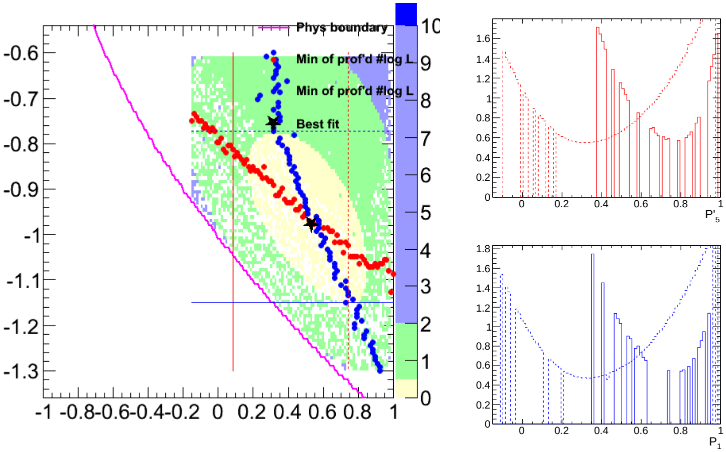
- 2D $\log\mathcal{L}$ for data
- $\star$ best fit
- magenta: physical limit
- red dots: min profiled $\log\mathcal{L}$ along P_1
- blue dots: likewise for P_5'

RIGHT

- profiled $\log\mathcal{L}$ for roughly $\log\mathcal{L}\sim 0.5$
- P_1 (red solid/dash)
- P_5' (blue solid/dash)

Bin 1



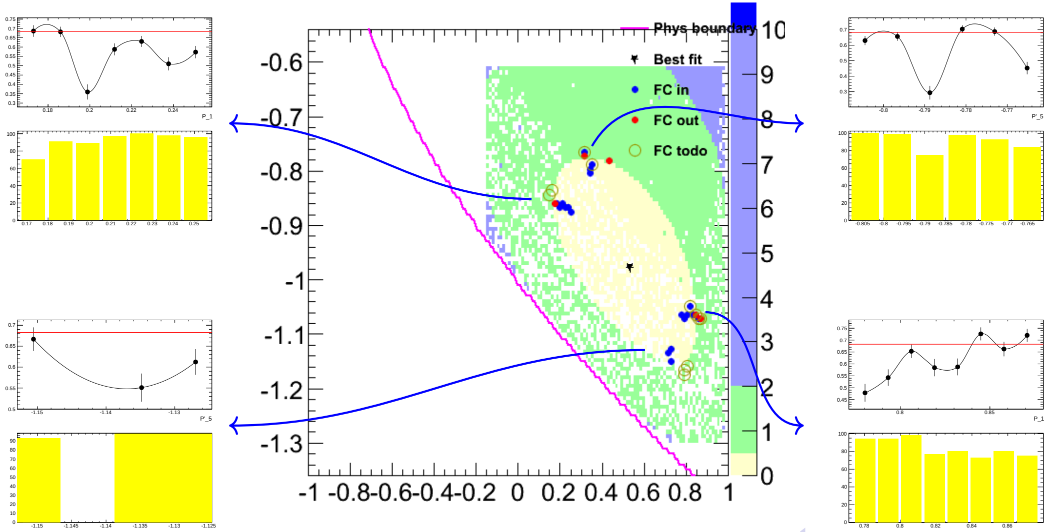


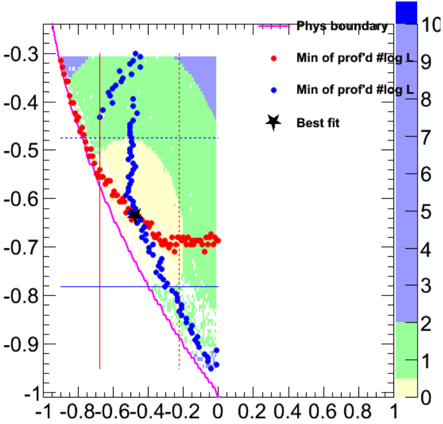
LEFT:

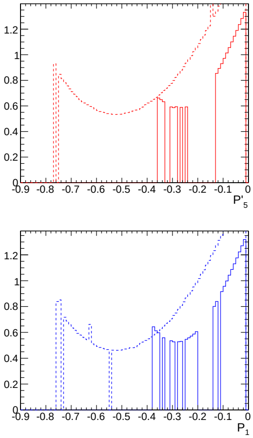
- 2D $\log \mathcal{L}$ for data
- $\star$ best fit
- magenta: physical limit
- red dots: min profiled $\log \mathcal{L}$ along P_1
- blue dots: likewise for P_5'

RIGHT

- profiled $\log \mathcal{L}$ for roughly $\log \mathcal{L} \sim 0.5$
- P_1 (red solid/dash)
- P_5' (blue solid/dash)


Phys boundary
Best fit
FC in
FC out
FC todo
P_1
P'_5



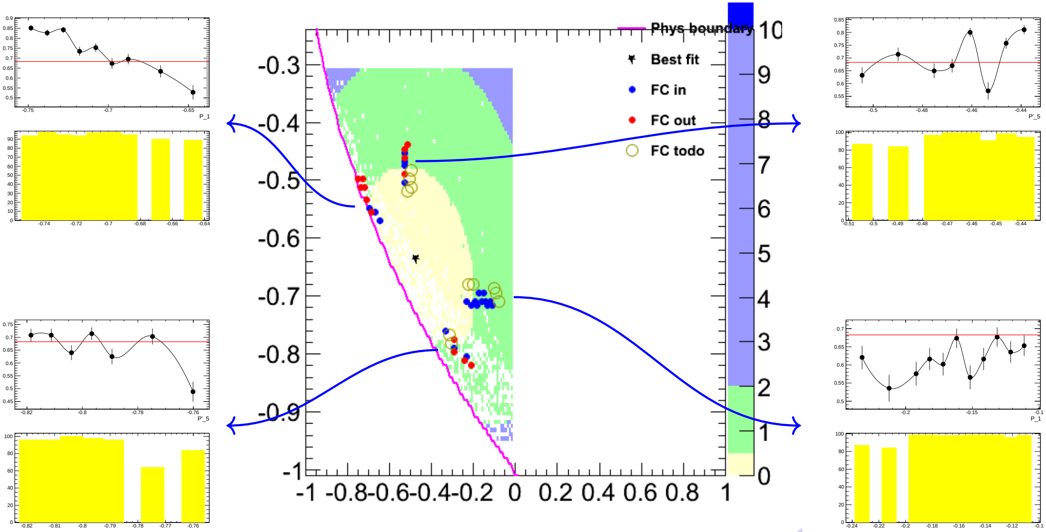


LEFT:

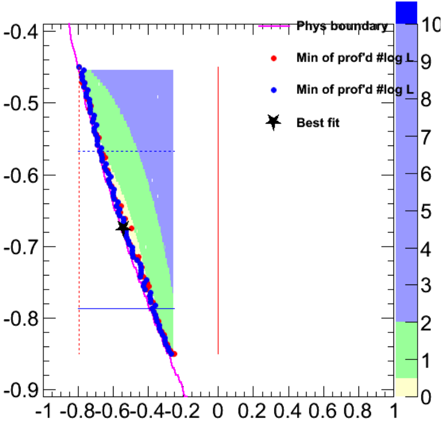
- 2D $\log \mathcal{L}$ for data
- $\star$ best fit
- magenta: physical limit
- red dots: min profiled $\log \mathcal{L}$ along P_1
- blue dots: likewise for P_5'

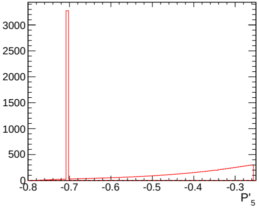
RIGHT

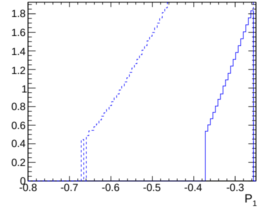
- profiled $\log \mathcal{L}$ for roughly $\log \mathcal{L} \sim 0.5$
- P_1 (red solid/dash)
- P_5' (blue solid/dash)


Phys boundary
Best fit
FC in
FC out
FC todo
P_1
P'_5

Bin 5





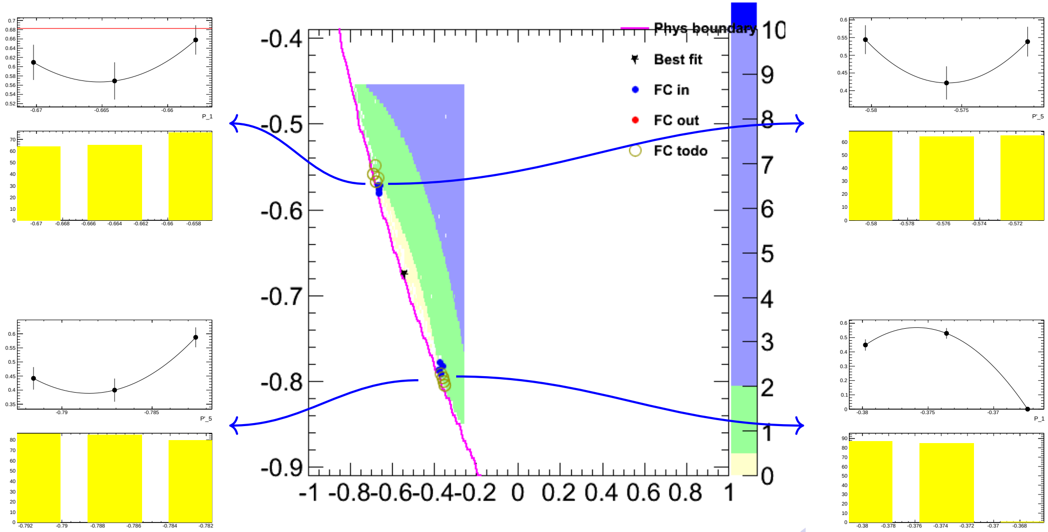


LEFT:

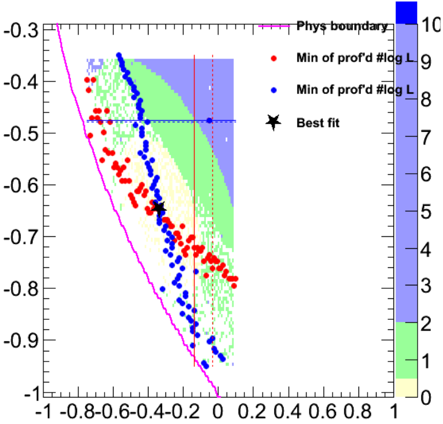
- 2D $\log \mathcal{L}$ for data
- $\star$ best fit
- magenta: physical limit
- red dots: min profiled $\log \mathcal{L}$ along P_1
- blue dots: likewise for P_5'

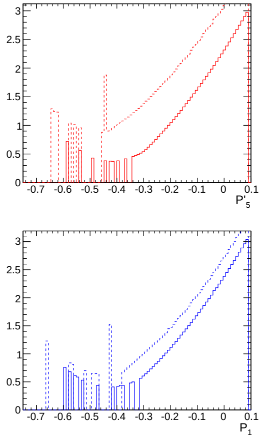
RIGHT

- profiled $\log \mathcal{L}$ for roughly $\log \mathcal{L} \sim 0.5$
- P_1 (red solid/dash)
- P_5' (blue solid/dash)


Phys boundary
Best fit
FC in
FC out
FC todo
P_1
P'_5
P'_5
P_1

Bin 7



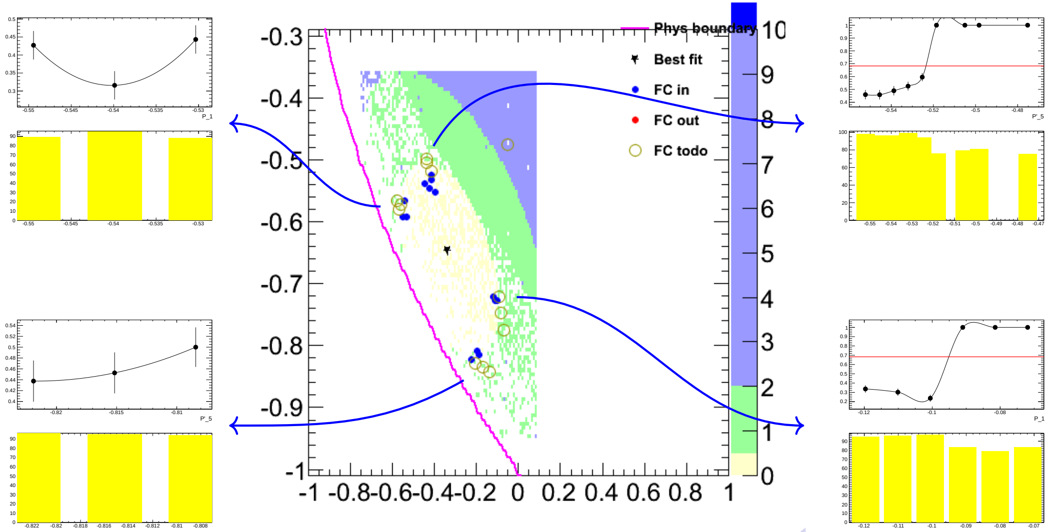


LEFT:

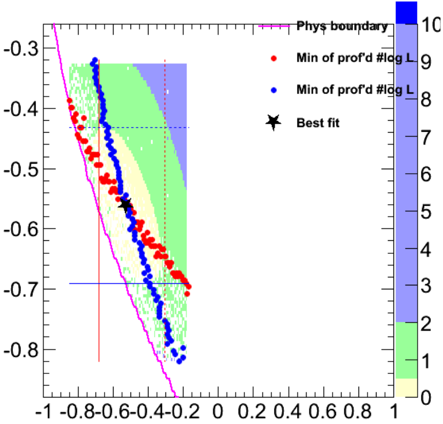
- 2D $\log\mathcal{L}$ for data
- $\star$ best fit
- magenta: physical limit
- red dots: min profiled $\log\mathcal{L}$ along P_1
- blue dots: likewise for P_5'

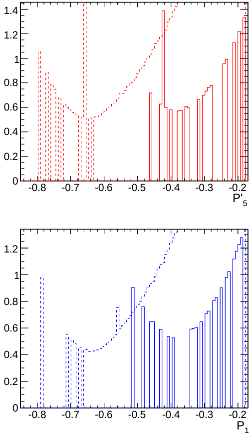
RIGHT

- profiled $\log\mathcal{L}$ for roughly $\log\mathcal{L} \sim 0.5$
- P_1 (red solid/dash)
- P_5' (blue solid/dash)



Phys boundary
Best fit
FC in
FC out
FC todo
P_1
P'_5



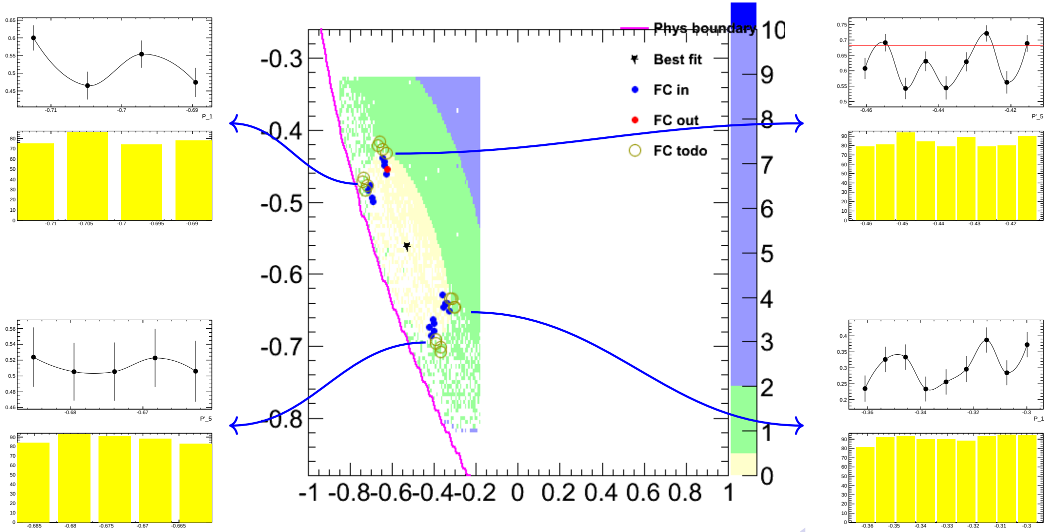


LEFT:

- 2D $\log \mathcal{L}$ for data
- $\star$ best fit
- magenta: physical limit
- red dots: min profiled $\log \mathcal{L}$ along P_1
- blue dots: likewise for P_5'

RIGHT

- profiled $\log \mathcal{L}$ for roughly $\log \mathcal{L} \sim 0.5$
- P_1 (red solid/dash)
- P_5' (blue solid/dash)


Phys boundary
Best fit
FC in
FC out
FC todo
P_1
P'_5
P'_5
P_1

Additional stuff

Additional or backup slides

Bibliography I