(Challenges of) Hadron Spectroscopy of Pentaquark Candidates at the LHCb experiment

Da Yu Tou
on behalf of the LHCb experiment

Tsinghua University

PWA13/ATHOS8, 28 May 2024

Contents

- Introduction
  - History
  - Motivation
- Post $\Lambda_b \to J/\psi p K$ results of the LHCb experiment:
  - $\Xi_b \to J/\psi \Lambda K$
  - $B_s \to J/\psi p \bar{p}$
  - $B \to J/\psi \Lambda \bar{p}$
- Ongoing amplitude analysis of $\Lambda_b \to J/\psi p K$ → working on this.

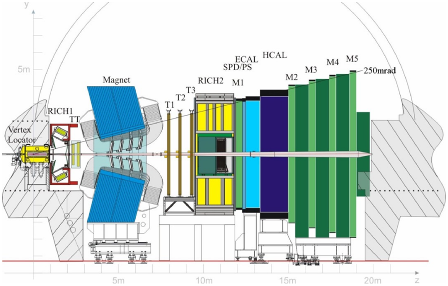
Disclaimer:

- Will focus primarily on challenges faced by amplitude analyses on pentaquark candidates at the LHCb experiment.
- There will be a heavy bias to $\Lambda_b \to J/\psi p K$ since I personally work on it.

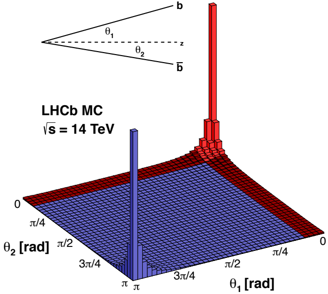
The LHCb Experiment

The LHCb Detector

- LHCb is a single-arm spectrometer in the forward region ($2 < \eta < 5$) at the LHC.
- Optimised for studying b (and c) decays which are boosted in the forward region.



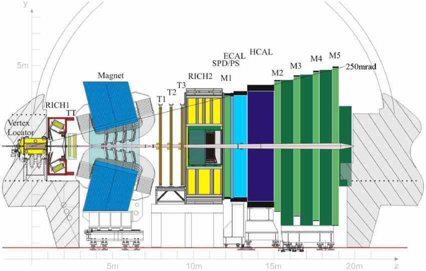
LHCb cross-section



Boosting of $b\bar{b}$ pair

The LHCb Detector

- Good tracking and particle identification performance.

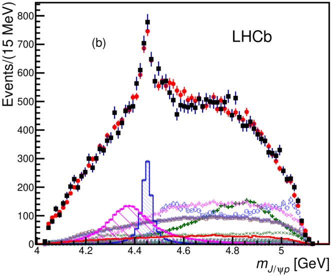


- VELO and T stations.
  - Momentum resolution $\sigma_p/p \approx 0.5 - 1\%$.
  - IP resolution $(15 + 29/p_T\ [\mathrm{GeV}])\mu m$.
- Muon, RICH, ECAL and HCAL.
  - $\epsilon(e) \sim 90\%$, $e \to h$ misID $\sim 5\%$.
  - $\epsilon(K) \sim 95\%$, $\pi \to K$ misID $\sim 5\%$.
  - $\epsilon(\mu) \sim 97\%$, $\pi \to \mu$ misID $\sim 1 - 3\%$.

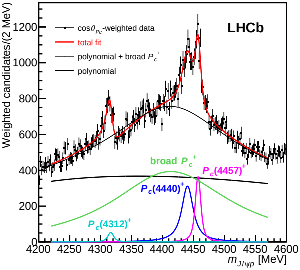
Introduction

$P_{c\bar{c}}^{+}$ Candidates Discovery - $\Lambda_b \to J/\psi pK$

- First discovered in $\Lambda_b \to J/\psi pK$ by LHCb in 2015.
  - $P_{c\bar{c}}(4380)^+ \to J/\psi p$
  - $P_{c\bar{c}}(4450)^+ \to J/\psi p$
- An updated mass spectrum analysis with full Run 1+2 data:
  - Found a narrow $P_{c\bar{c}}(4312)^+$
  - Resolved $P_{c\bar{c}}(4450)^+$ into narrow $P_{c\bar{c}}(4440)^+$ and $P_{c\bar{c}}(4457)^+$



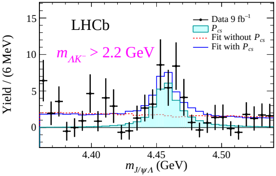
PRL 115 (2015) 072001



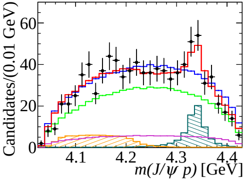
PRL 122 (2019) 222001

More $P_{c\bar{c}}^{+}$ and $P_{c\bar{c}s}^{0}$ Joins the Party

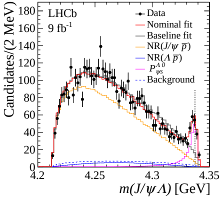
- The discoveries in $\Lambda_b \to J/\psi pK$ channel led to searches in other channels.
- Successful searches so far at the LHCb experiment:
  1. $P_{c\bar{c}s}(4459)^0 \to J/\psi\Lambda$ in $\Xi_b \to J/\psi\Lambda K$
  2. $P_{c\bar{c}}(4337)^{\pm} \to J\psi p^{\pm}$ in $B_s \to J/\psi p\bar{p}$
  3. $P_{c\bar{c}s}(4338)^0 \to J/\psi\Lambda$ in $B \to J/\psi\Lambda\bar{p}$



Sci.Bull. 66 (2021) 1278-1287



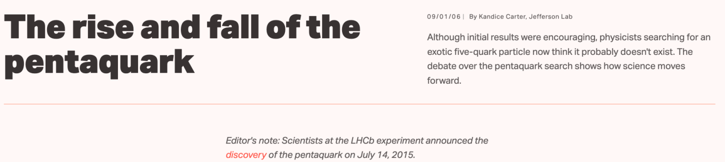
PRL 128 (2022) 062001



PRL 131 (2023) 031901

Historical Pentaquarks

- Predicted back in 1987 by PLB 195 (1987) 484-488 and PLB 193 (1987) 323.
- Rather interesting reading PDG 2006 and PDG 2008 review of early pentaquarks.
    - ▶ $\Theta(1540)^+$, $\Phi(1860)$ and $\Theta_c(3100)^+$.
    - ▶ Failed to be replicated, PDG 2008 excerpt: "There are two or three recent experiments that find weak evidence for signals (pentaquarks) near the nominal masses, but there is simply no point in tabulating them in view of the overwhelming evidence that the claimed pentaquarks do not exist."



The rise and fall of the pentaquark

09/01/06 | By Kandice Carter, Jefferson Lab

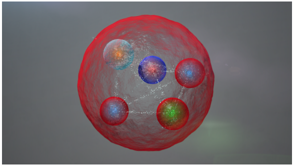
Although initial results were encouraging, physicists searching for an exotic five-quark particle now think it probably doesn't exist. The debate over the pentaquark search shows how science moves forward.

Editor's note: Scientists at the LHCb experiment announced the discovery of the pentaquark on July 14, 2015.

September 2006 Symmetry Article

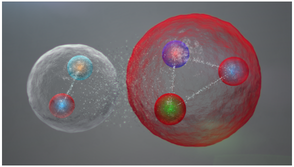
Physics Motivation

- Studies are ongoing to understand the pentaquark candidates.
- Some interpretations:



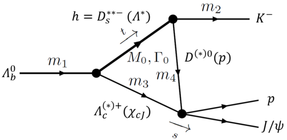
Compact Pentaquark

JHEP 12 (2015) 128
PLB 749 (2015) 289-291
PRD 95, 054027
PLB 793 (2019) 365-371



Baryon-Meson Molecule

PRL 115 (2015) 122001
PLB 753 (2016) 547-551
PRC 85 (2012) 044002



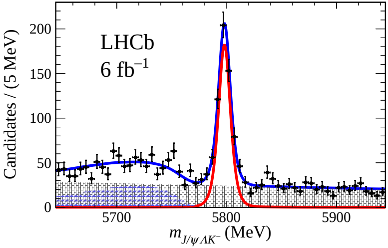
Triangle Diagram

PRD 92 (2015) 071502

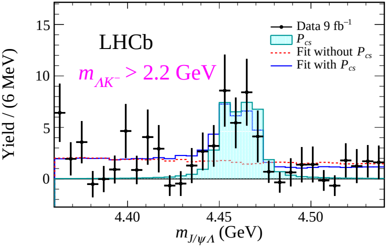
Post $\Lambda_b \to J/\psi pK$ Discoveries

$P_{c\bar{c}s}(4459)^0$ in $\Xi_b \to J/\psi\Lambda K$

- 1750 ± 50 signal candidates.[Sci. Bull. 66 (2021) 1278-1287]
- Adding a $P_{c\bar{c}s}(4459)^0$ improves $\Xi \to \Lambda K$ only amplitude fit.
  - 3.1σ after accounting for systematic uncertainties.
- $P_{c\bar{c}s}(4459)^0$ peaks 19 MeV below $\Xi_c^0\bar{D}^{*0}$ mass threshold.
  - PRD 101, 034018 predicts two states ($J^P = \frac{1}{2}^-, \frac{3}{2}^-$) here.
  - Insufficient statistics to probe this hypothesis.



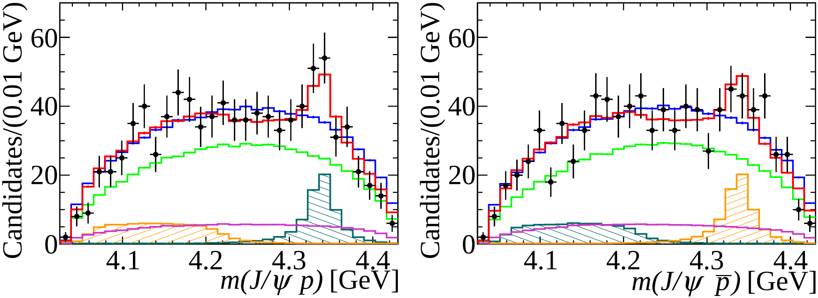
Ξ_b^- mass fit.



Two $P_{c\bar{c}s}^0$ fit.

$P_{c\bar{c}}(4337)^+$ in $B_s \to J/\psi p\bar{p}$

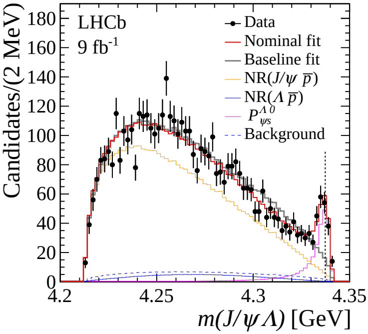
- 797 ± 31 signal candidates. [PRL 128 (2022) 062001]
- Helicity formalism found $P_{c\bar{c}}(4337)^+$ with 3.1σ but due to limited statistics:
  - Equal $P_{c\bar{c}}^+$ and $P_{c\bar{c}}^-$ coupling with a phase difference.
  - Cannot distinguish between four J^P hypothesis: $\frac{1}{2}^\pm, \frac{3}{2}^\pm$.

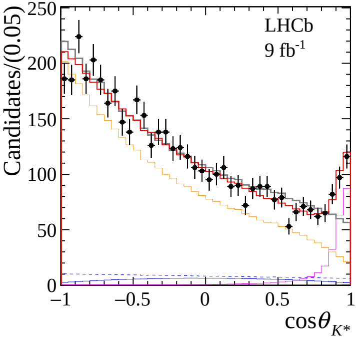


$m(J/\psi p)$ $m(J/\psi\bar{p})$

$P_{c\bar{c}s}(4338)^0$ in $B \to J/\psi\Lambda\bar{p}$

- About 4400 signal candidates. [PRL 131 (2023) 031901]
- Small Q-value of 128 MeV. Observed $P_{c\bar{c}s}(4338)^0$ near $\Xi_c^+ D^-$ threshold.
    - $J^P = \frac{1}{2}^-$ preferred, $J^P = \frac{1}{2}^+$ excluded at 90% confidence.
- Tried looking for $P_{c\bar{c}s}(4255)^0$ (near $\Lambda_c^+ D_s^-$ threshold) but it is statistically insignificant.





Low Statistics

- Statistics of these 3 channels are quite low. (1750, 797 and 4400)
- They are barely able to find an evidence or observation but:
  - Cannot probe two-peak hypothesis for $P_{c\bar{c}s}(4459)^0$.
  - Cannot determine J^P of $P_{c\bar{c}}(4337)^+$.
  - $P_{c\bar{c}s}(4255)^0$ not statistically significant.
- Look out for results by Run 3 detector!
  - Higher luminosity → more $b\bar{b}$.
  - More efficient software trigger → do fast reconstruction and trigger on tracks/vertices rather than simple hardware triggers.

Back to $\Lambda_b \rightarrow J/\psi p K$

Run 1+2 Update of $\Lambda_b \to J/\psi pK$

- What will we improve on?
- Implications from mass spectrum analysis:
  1. Statistics Run 2 will give us access to narrow $P_{c\bar{c}}(4312)^+$ and resolve $P_{c\bar{c}}(4440)^+$ -$P_{c\bar{c}}(4457)^+$ pair.
  2. Consider mass resolution.
  3. Consider coupled channel effect.
- Similar helicity model as Run 1 data.

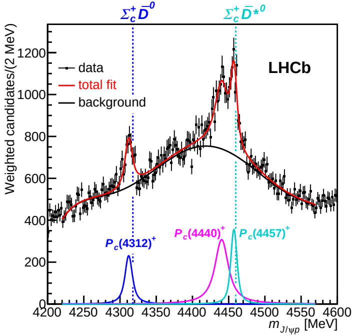
Implications of Mass Spectrum Analysis

- PRL 122 (2019) 222001 showed that statistics in Run 2 will give us access to narrow $P_{c\bar{c}}(4312)^+$ and resolve $P_{c\bar{c}}(4440)^+$ -$P_{c\bar{c}}(4457)^+$ pair.
- Widths of narrow $P_{c\bar{c}}^+$ candidates is very close to $m(J/\psi p)$ resolution ($\Delta m \sim 2.5$ MeV).
  - Use same selection and mass resolution model as mass spectrum paper.
  - Numerically integrate mass resolution in amplitude fit.

State	M [MeV]	Γ [MeV]
$P_{c\bar{c}}(4312)^+$	$4311.9 \pm 0.7^{+6.8}_{-0.6}$	$9.8 \pm 2.7^{+3.7}_{-4.5}$
$P_{c\bar{c}}(4440)^+$	$4440.3 \pm 1.3^{+4.1}_{-4.7}$	$20.6 \pm 4.9^{+8.7}_{-10.1}$
$P_{c\bar{c}}(4457)^+$	$4457.3 \pm 0.6^{+4.1}_{-1.7}$	$6.4 \pm 2.0^{+5.7}_{-1.9}$

Implications of Mass Spectrum Analysis

- PRL 122 (2019) 222001: Masses of narrow $P_{c\bar{c}}^{+}$ candidates slightly below open charmed baryon-meson mass thresholds.
  - Coupled channel effects need to be considered via $\mathcal{K}-$matrix parameterization of $m(J/\psi p)$ system.
  - E.g. $f_0(980) \to \pi\pi$ cusp due to strong coupling to KK. PLB 63 (1976) 224-227

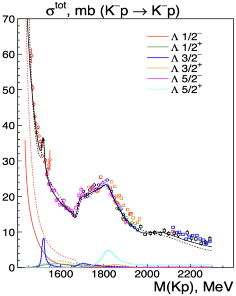


Challenges

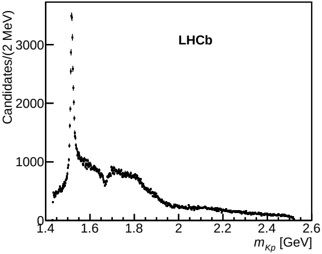
- Lack of Λ^* spectroscopy data above 2 GeV.
  - Lack of good constraints on what resonances live here.
  - Need to fit new states.
- Small contribution of (signal) $P_{c\bar{c}}^+$ states.
  - Likelihood is dominated by Λ^* system (background/nuisance).
- High amount of statistics.
  - Need an (effective) model with very good description of data.
  - Long fit time.
  - Start finding numerical issues.

High $m(Kp)$

- PDG has a lot of constraints for Λ^* states below 2 GeV.
- Not many constraints of Λ^* spectroscopy above 2 GeV - only 3 states with 3/4 stars in PDG.
  - EPJA 55 (2019) 10, 179 has a review showing lack good data for Kp scattering above 2 GeV. (left plot)
- Needed to add Λ^* states to get a good description of data.



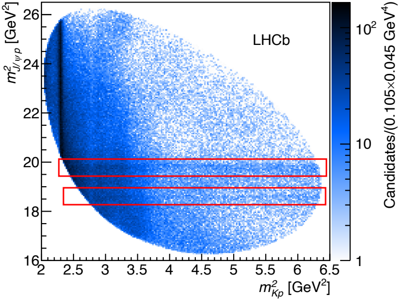
Kp scattering data



$m(Kp)$ spectrum in $\Lambda_b \to J/\psi pK$

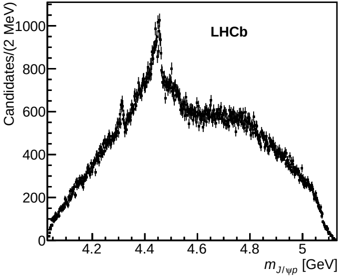
High $m(Kp)$

- This makes life difficult for $P_{c\bar{c}}^{+}$ amplitude analysis.
  - Looking at $\Lambda_b \to J/\psi pK$ Dalitz plot below, this is the region where $P_{c\bar{c}}^{+}$ contribution is very dominant.
- Isobar model predicts tons of Λ^* in this region.
- Needed to add a few new states to get good description of data.
- The fitter likes to make them broad ($\Gamma = 500$ MeV) - probably absorbs the multiple Λ^* into one broad state.



$m(J/\psi p)$ $\mathcal{K}$ – matrix

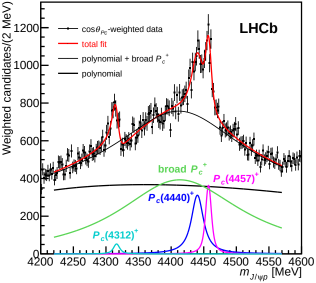
- Non-resonant components of $m(J/\psi p)$ $\mathcal{K}$ – matrix.
  - Large fit fraction $\mathcal{O}(5-10\%)$.
  - Strongly interferes with Λ^* system.
  - Suspect fitter is compensating to get a good effective description of Λ^* $/m(Kp)$ system.



Full $m(J/\psi p)$ spectrum in $\Lambda_b \to J/\psi pK$. [PRL 122 (2019) 222001]

$m(J/\psi p)$ $\mathcal{K}$ – matrix

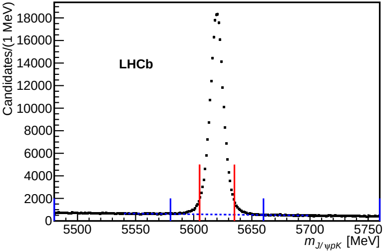
- Small contribution of $P_{c\bar{c}}^{+}$ states.
  - Fit fraction $\mathcal{O}(1\%)$: likelihood is dominated by the Λ^{*} states and non-resonant $m(J/\psi p)$.
  - Sometimes the likelihood minimization cannot resolve two-peak structure around 4500 MeV.
  - Solution: run multiple amplitude fits with randomized initial parameters.



Two peak structure in $m(J/\psi p)$ spectrum, $m(Kp) > 1.9$ GeV.
PRL 122 (2019) 222001

High Statistics

- Amplitude fits on high statistics data demand a very good description of data. Same selection as PRL 122 (2019) 222001.
  - $\sim$ 240k signal.
  - $\sim$ 140k background.



Model for High Statistics

- $m(Kp)$ lineshape consists of 2 splines and Breit-Wigners.
- $m(J/\psi p)$ $\mathcal{K}-$matrix for different J^P.
- End up with 389 free parameter model.
    - Many of these parameters are correlated.
    - Covariance matrix is not positive definite.
    - Good physics description but terrible statistics/minimization model.

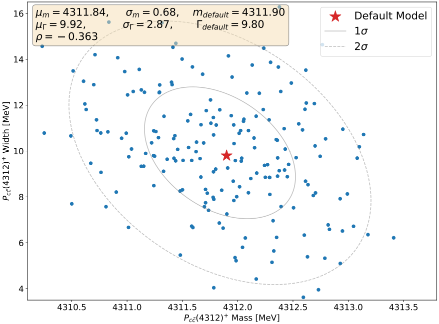
	m_0	$\sigma(m_0)$	Γ_0	$\sigma(\Gamma_0)$	J^P
$\Lambda^*(1600)$	1592	10	150	28	$\frac{1}{2}^+$
$\Lambda^*(1710)$	1713	13	180	42	$\frac{1}{2}^+$
$\Lambda^*(1810)$	1821	10	174	50	$\frac{1}{2}^+$
$\Lambda^*(1890)$	1900	5	161	15	$\frac{3}{2}^+$
$\Lambda^*(1520)$	1519.6	0.5	17	1	$\frac{3}{2}^-$
$\Lambda^*(1690)$	1691	3	54	5	$\frac{3}{2}^-$
$\Lambda^*(2050)$	2056	22	493	61	$\frac{3}{2}^-$
$\Lambda^*(1830)$	1820	4	114	10	$\frac{5}{2}^-$
$\Lambda^*(2580)$	2580	-	250	-	$\frac{5}{2}^-$
$\Lambda^*(1820)$	1823.5	0.8	89	2	$\frac{5}{2}^+$
$\Lambda^*(2110)$	2036	13	200	38	$\frac{5}{2}^+$
$\Lambda^*(2020)$	2043	22	200	75	$\frac{7}{2}^+$
$\Lambda^*(2100)$	2086	6	305	16	$\frac{7}{2}^-$
$\Lambda^*(2350)$	2350	15	150	15	$\frac{9}{2}^+$

[PRC 88 (2013) 035205]

Covariance Matrix

- Fits can converge, but covariance matrix is never positive-definite.
- No longer calculate covariance matrix - avoid (massive) $\mathcal{O}(N^2_{\text{parameter}} \times N_{\text{candidates}})$ computing cost.
- Statistical uncertainties will be evaluated with toys/pseudoexperiments.

Example fit to toy results for statistical uncertainties ($P_{c\bar{c}}(4312)^+$ mass and width from PRL 122 (2019) 222001).



Computation

- The negative log-likelihood ($-\log(\mathcal{L})$) calculation is done on GPUs written in `CUDA`.
- This is interfaced with RooFit (and then to `MINUIT` as the minimizer).
  - Use numerical gradients calculated by finite difference.
- Running on GPUs is necessary for speed.
  - But still takes 0.5-1.5 days on a single RTX 3090.
  - Syracuse University was kind enough to give us access to OrangeGrid, $\sim$ 100 GPUs.

- Crucial to run a lot of fits:
  - Statistical uncertainties - toys.
  - Best default models - randomized initial parameters fit.

Numerical

- For a long time it was very hard to get fits to converge.
  - Half of fits were failing.
  - Most of systematic uncertainties fit were not converging.
- Debugging was quite a learning experience.

Numerical

- `MINUIT` internally uses a quasi-Newton method to minimize.*
- It needs first derivatives to figure out where to go next.
- TensorWaves, tf-pwa and TFA2 uses automatic differentiation libraries (`Tensorflow`, `pytorch`, `JAX`) to do this.
- When analytical integral is unavailable to `MINUIT`, it uses finite difference instead.

$$\frac{\partial - \log(\mathcal{L})}{\partial x_i} = -\frac{1}{\mathcal{L}}\frac{\partial \mathcal{L}}{\partial x_i}, \quad \frac{\partial \mathcal{L}}{\partial x_i} = \frac{\mathcal{L}(x_i + \delta) - \mathcal{L}(x_i - \delta)}{2\delta}$$

---

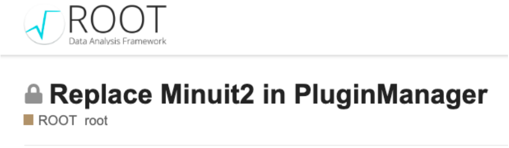
*`MINUIT` uses Davidon–Fletcher–Powell formula. Davidon broke into an FBI office and stole documents which led to an inverstigation and reforms of FBI.

Numerical

- Numerically inaccurate computations can lead to badly behaving fits.
    - Bad estimate for first derivatives.
- First thing is to check if we were losing floating-point precision summing per-event $-\log(\mathcal{L})$ over 400k events.
- Kahan sum (floating point compensated) vs simple sum (what we were using) difference is orders of magnitude smaller than $\Delta - \log(\mathcal{L})$ used in finite difference.
- Summation is not the problem.

Hacking ROOT

- MINUIT2 object does not have a public function that provides first derivatives.
- But it has a private member that keeps a history of all your finite difference $\Delta - \log(\mathcal{L})$, δ and first derivatives.
- Not super hard to hack this thanks to ROOT developer help:
  - Redefine MINUIT3 object, everything the same except a public function that exposes the derivative history.
  - After adding ROOT runtime library loading magic, you can get all the first derivatives in your amplitude fit.



Replace Minuit2 in PluginManager

ROOT root

Link Asking ROOT developers for help with ROOT runtime library.

$\frac{\partial \mathcal{L}}{\partial x}$ Debugging

$$\mathcal{A}_c = \sum_b \left[1 - i\mathcal{K}\rho n^2\right]^{-1}_{cb} \mathcal{P}_b$$

- $m(J/\psi p)$ $\mathcal{K}-$ matrix parameters have large first derivatives.
  - Sign of numerical instability.
- Mikhail Mikashenko suggested reparameterizing m_R in $\mathcal{K}-$ matrix to decorrelated $\mathcal{K}-$ matrix m_R pole and g coupling.
  - $\mathcal{K}$-matrix: $\mathcal{K}_{ab} = \sum_R \frac{g_a^R g_b^R}{M_R^2 - m^2} + b_{ab}$
  - Old: $m_R = m_{\mathrm{float}}$
  - New: $m_R^2 = m_{\mathrm{float}}^2 + \mathcal{R}\left[i g_c^2 \rho_c(m_{\mathrm{float}})\right]$
- Somewhat helps, but fits still unstable and often go into a region with `NaN` $-\log(\mathcal{L})$.
  - Mikhail said `NaN` happens due to matrix inversion in $\mathcal{K}-$ matrix formula.

Matrix Inversion

- We were inverting $N \times N$ complex matrix by splitting into real and imaginary, then invert each individually

$$\begin{aligned} \text{let} \quad V &= 1 - i\mathcal{K}\rho \\ V &= A + iB \\ \text{and} \quad T &= V^{-1} \\ T &= C + iD \end{aligned} \quad \rightarrow \quad \begin{aligned} VT &= \mathbb{I} + i\mathbf{0} \\ AC - BD &= \mathbb{I} \\ AD + BC &= \mathbf{0} \\ \text{gives} \qquad C &= (A - BA^{-1}A)^{-1} \\ D &= -(B + AB^{-1}A)^{-1} \end{aligned}$$

- Mathematically valid but:
  - ✗ Both A and B needs to be invertible (otherwise `NaN`).
  - ✗ Many computation steps → lose precision.
  - ✗ Only need first row of $(1 - i\mathcal{K}\rho)^{-1}$ ($J/\psi p$ channel lineshape).

Non-physics Challenges - Matrix Solver

- $LU^{\dagger}$ decomposition and solve linear equation‡ by forward-backward substitution. E.g. 3 channels:

$$\left[1 - i\mathcal{K}\rho n^2\right] \begin{pmatrix} \mathcal{A}_c \\ 0 \\ 0 \end{pmatrix} = \mathcal{P} \quad \xrightarrow{\text{solve}} \quad \begin{array}{l} \text{forward sub: } Ly = \mathcal{P} \\ \text{backward sub: } U \begin{pmatrix} \mathcal{A}_c \\ 0 \\ 0 \end{pmatrix} = y \end{array}$$

$$LU \text{ decomp:} \quad \left[1 - i\mathcal{K}\rho n^2\right] = LU$$

- LU decomposition of 4×4 matrix:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ L_{21} & 1 & 0 & 0 \\ L_{31} & L_{32} & 1 & 0 \\ L_{41} & L_{42} & L_{43} & 1 \end{pmatrix} \begin{pmatrix} U_{11} & U_{12} & U_{13} & U_{14} \\ 0 & U_{22} & U_{23} & U_{24} \\ 0 & 0 & U_{33} & U_{34} \\ 0 & 0 & 0 & U_{44} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix}$$

[†] L = lower triangular matrix, U = upper triangular matrix

[‡] As wise man on a blog once said: "Don't invert it, solve it!"

Aftermath

- Most minimization fit ($\sim 90\%$) now converging, including:
  - Randomized initial parameters fit for best default model.
  - Toys for systematic uncertainties.
- Status of analysis:
  - Very close to completion and a collaboration review.
  - A few systematics left to figure out.

Summary

Human Resource Challenge

- The PhD/postdoc churn: initial analysts graduated and left to other institutes as postdoc or found jobs in private sector.
- I started working on this about a year and a half ago because an ex-student has to focus on the priorities of his postdoc institute.
- Need overhead period of a few months to grasp what is going on.
  - Amplitude analysis is quite complex.
  - Also needed to understand what is going on inside the code.

Summary

- LHCb results of pentaquarks:
  - Found 3 $P^{+}_{c\bar{c}}$ candidates and 1 $P^{0}_{c\bar{c}s}$ candidate.
  - Has evidence for 1 $P^{+}_{c\bar{c}}$ candidate and 1 $P^{0}_{c\bar{c}s}$ candidate.
- $\Xi_b \to J/\psi \Lambda K$, $B_s \to J/\psi p\bar{p}$ and $B \to J/\psi \Lambda \bar{p}$ had to deal with low statistics.
- $\Lambda_b \to J/\psi pK$ have high statistics which exposes a host of other problems:
  - Minimization demands a very good description of data.
  - Consume a lot of computational resources.
  - Large statistics will also expose numerical problems.

Back Up

Updated Helicity Formalism $\Lambda_b \to J/\psi p K$

1. Tune decay angle definitions. [Chin.Phys.C 45 (2021) 063103]
2. Added particle two factor. [PRD 101 (2020) 034033]