Amplitude analysis of four-body decays using a massively-parallel fitting framework

To cite this article: C Hasse *et al* 2017 *J. Phys.: Conf. Ser.* **898** 072016

IOP Conf. Series: Journal of Physics: Conf. Series **898** (2017) 072016 doi:10.1088/1742-6596/898/7/072016

Amplitude analysis of four-body decays using a massively-parallel fitting framework

C Hasse[1,2,3], J Albrecht[2], A A Alves Jr.[3], P d'Argent[4], T D Evans[5], J Rademacker[6] and M D Sokoloff[3]

[1] CERN, CH-1211 Geneva 23, Switzerland
[2] Experimentelle Physik V, Technische Universität Dortmund, Otto-Hahn-Straße 4, 44227 Dortmund, Germany
[3] Physics Department, University of Cincinnati, 2600 Clifton Ave. Cincinnati, OH 45221, USA
[4] Physikalisches Institut, Universität Heidelberg, Im Neuenheimer Feld 226, 69120 Heidelberg, Germany
[5] Department of Physics, University of Oxford, Parks Road, Oxford OX1 3PU, UK
[6] School of Physics, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK

E-mail: christoph.hasse@cern.ch, johannes.albrecht@cern.ch, antonio.augusto.alves.junior@cern.ch, p.dargent@cern.ch, timothy.david.evans@cern.ch, jonas.rademacker@bristol.ac.uk and sokoloff@ucmail.uc.edu

Abstract. The GooFit Framework is designed to perform maximum-likelihood fits for arbitrary functions on various parallel back ends, for example a GPU. We present an extension to GooFit which adds the functionality to perform time-dependent amplitude analyses of pseudoscalar mesons decaying into four pseudoscalar final states. Benchmarks of this functionality show a significant performance increase when utilizing a GPU compared to a CPU. Furthermore, this extension is employed to study the sensitivity on the $\mathrm{D}^0 - \overline{\mathrm{D}}^0$ mixing parameters x and y in a time-dependent amplitude analysis of the decay $\mathrm{D}^0 \to \mathrm{K}^+\pi^-\pi^+\pi^-$. Studying a sample of 50 000 events and setting the central values to the world average of $x = (0.49 \pm 0.15)\,\%$ and $y = (0.61 \pm 0.08)\,\%$, the statistical sensitivities of x and y are determined to be $\sigma(x) = 0.019\,\%$ and $\sigma(y) = 0.019\,\%$.

1. Introduction

In physics analyses it is common to fit a theoretical model to observed data to extract parameters of interest. This involves minimizing the differences between a model and data, which is mostly done by performing a minimization of a cost function, for example the likelihood function. However, problems arise because the computations become very expensive as the complexity of the models and number of events increases. The GooFit [1–3] framework has been designed to address this issue by allowing such computations to be performed in parallel. It is built upon the Thrust library [4] to be able to run on different parallel architectures, while maintaining a control flow similar to the RooFit package [5], which is commonly used in high energy physics to fit theoretical models to data, and which only runs on CPUs. While GooFit has been successfully employed in several analyses, even for complex models such as time-dependent mixing in three-body decays, it did not allow for performing a time-dependent amplitude analyses of four-body decays. This functionality was recently added and will be described in this paper.

2. Mixing in the decay $D^0 \to K^+\pi^-\pi^+\pi^-$

Mixing or oscillation of neutral mesons is a process during which a particle transitions into its antiparticle or vice versa. This process has been observed in the K^0, B^0, B_s^0 and D^0 systems. The D^0 system is the only one comprised of up-type quarks. One possible decay to study the

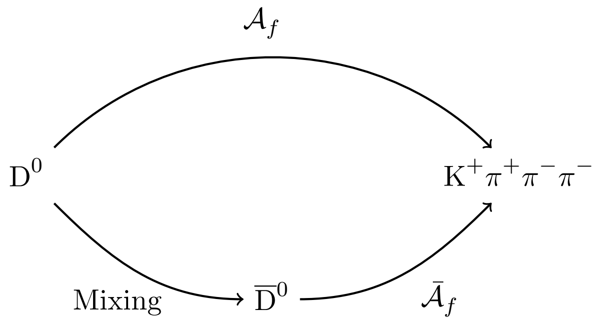

Figure 1. Schematic view of the two possible decay paths for a D^0 decaying into a K^+ π^- π^+ π^- final state. The top path corresponds to the direct decay, while the bottom path shows the mixing transition of a D^0 into a $\overline{D}^0$ followed by a decay into the final state.

phenomenon of mixing in the neutral charm meson system is the decay of D^0 to $K^+\pi^-\pi^+\pi^-$. This decay can proceed via two different decay amplitudes, which are depicted in figure 1. The top arrow depicts the direct decay subscribed $\mathcal{A}_f$, while the bottom arrow represents the decay proceeding via mixing into a $\overline{D}^0$ which decays into the final state via an amplitude subscribed $\bar{\mathcal{A}}_f$. Due to the mixing of a D^0 into a $\overline{D}^0$ being time-dependent, the overall decay rate becomes time-dependent. Analysing such time-dependent decay rates allows extraction of mixing properties of the D^0 system.The expression for the time-dependent decay rate of the D^0, assuming no CP violation, can be derived to be [6],

$$\begin{aligned}\frac{d\Gamma(\mathcal{A}_f)}{e^{-\Gamma t}\mathcal{N}_f} &= \left(\left|\mathcal{A}_f\right|^2 + \left|\bar{\mathcal{A}}_f\right|^2\right)\cosh(y\Gamma t) + \left(\left|\mathcal{A}_f\right|^2 - \left|\bar{\mathcal{A}}_f\right|^2\right)\cos(x\Gamma t) \\ &\quad -2\Re\left(\mathcal{A}_f\bar{\mathcal{A}}_f^*\right)\sinh(y\Gamma t) - 2\Im\left(\mathcal{A}_f\bar{\mathcal{A}}_f^*\right)\sin(x\Gamma t)\,.\end{aligned} \quad (1)$$

Most of the complexity of this expression lies within the model used to describe the two amplitudes $\mathcal{A}_f$ and $\bar{\mathcal{A}}_f$

3. Structure and implementation of four-body amplitudes

While equation (1) is completely general, the amplitudes that encode the properties of the decay are functions of the position in phase space occupied of the final state of the decay. The amplitude structure of a four-body decay is significantly more complicated than that of three-body decays because their phase space is five dimensional while three-body decays merely occupy a two-dimensional phase space.

Similar to other amplitude models, the implemented functionality assumes that multi-body decays mostly proceed via quasi two-body processes, which include two-body resonances. This leads to two possible decay chain topologies depicted in figure 2, where R_1 and R_2 are intermediate resonances and a, b, c and d are the four final decay products, in various configurations. Here, R_1 and R_2 can take the form of multiple kinematically allowed resonance states, resulting in many possible decay chains. A complete amplitude will therefore be modelled

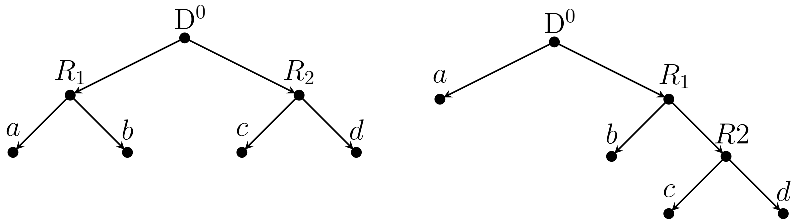

Figure 2. Possible quasi two-body decay topologies of a four-body decay. Left, a D^0 meson decays into two resonances R_1 and R_2 , which decay into two particles each. Right, a D^0 meson decays into a particle a and a resonance R_1 , which proceeds to decay into a resonance R_2 and a final state particle. R_2 then decays into the remaining two final state particles.

by a coherent sum over these decay chains $\mathcal{A}_i$ as,

$$\mathcal{A}_f = \sum_i c_i \mathcal{A}_i \,, \qquad c_i, \mathcal{A}_i \in \mathbb{C}. \tag{2}$$

Each decay chain $\mathcal{A}_i$ is constructed by the user from classes representing form factors, spin factors, resonance lineshapes, and possibly, in the case of two identical final state particles, Bose-symmetrization. After successfully constructing all necessary decay chains the user constructs two amplitude class instances representing $\mathcal{A}_f$ and $\bar{\mathcal{A}}_f$, which each hold the necessary decay chains to fit the theoretical model. The model creation is finalized by creating an instance of the time-dependent amplitude model class and passing the two amplitudes just created by the user. Upon creation the time-dependent model class automatically checks for recurring form factors, spin factors, and lineshapes in all decay chains. In case of multiple occurrences, these instances are substituted by a link to a single instance, thus removing redundant calculations. The proceeding steps of the internal model building process are explained in detail in [1, 2].

3.1. Normalization and event generation
During the fitting procedure the complete expression in equation (1) must be normalized accurately. As it is not feasible to find an analytic expression for such a complex function, the normalization is computed numerically. In our study, this requires evaluating the function at several million phase space points. To achieve a sufficiently fast generation of phase space events, we integrated the MCBooster library [7, 8], which allows very fast generation of phase space events on the GPU. This also enables the generation of pseudo-events, which are uniformly distributed phase space events weighted by the previously created amplitude model.

3.2. Validation
As this work implemented various new building blocks to model four-body decay amplitudes in GooFit it was important to validate the correctness of each of these new components. A cross check of the implementation was performed by comparing the newly implemented functionality of GooFit to the software package MINT3 [9]. MINT3 is based upon the MINT (Minuit Interface) package [10], which is used to perform time-integrated amplitude analyses of three- and four-body decays. Additionally, it supports the generation of pseudo-events. We generate $500,000$ pseudo-events for a specific amplitude model, which includes all newly implemented building blocks, and compare the resulting event samples. This comparison is performed by studying the phase space projections of the samples given the five variables $m_{12}, m_{34}, \cos_{12}, \cos_{34}$ and ϕ, where

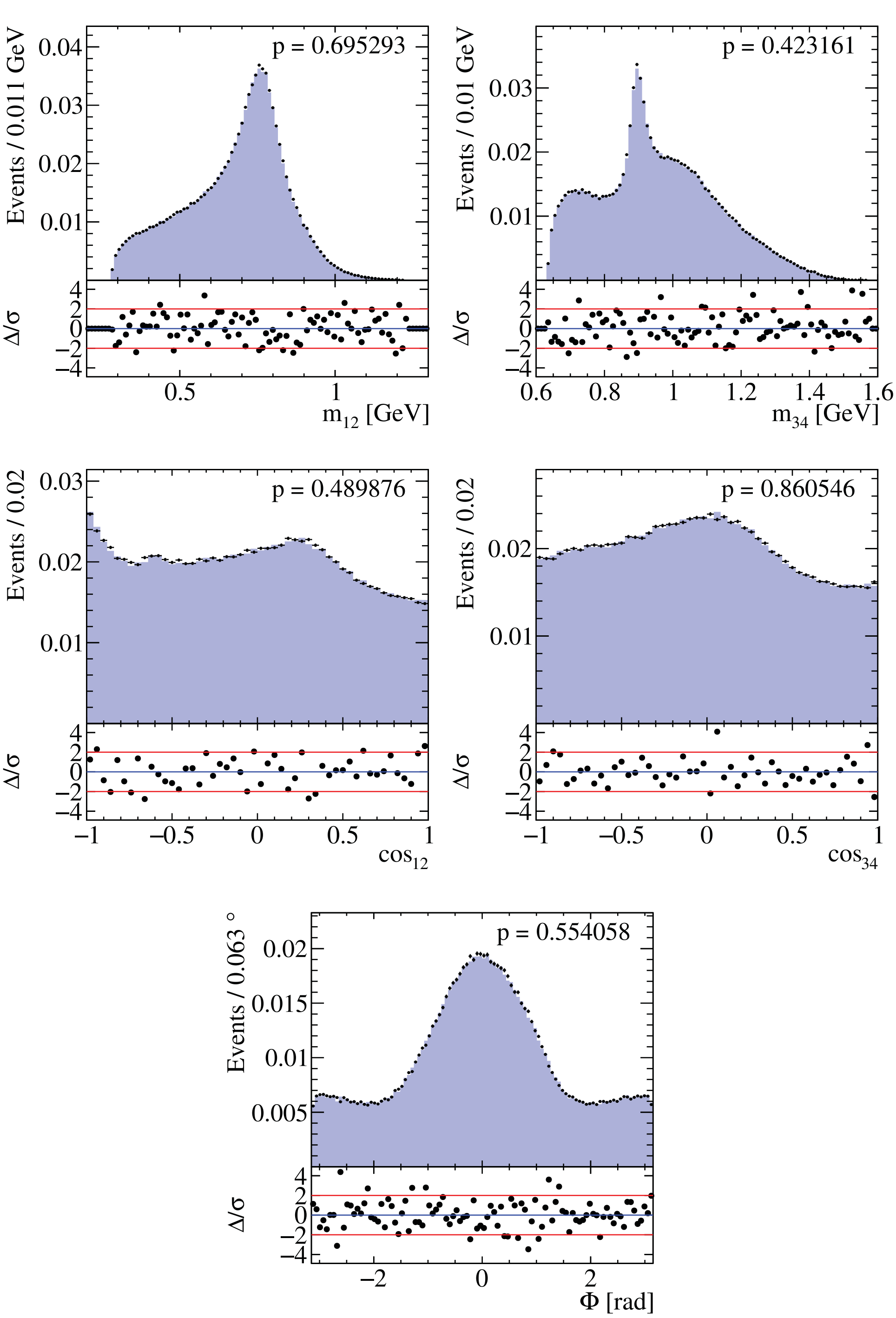

Figure 3. Comparison between the generated pseudo events from the MINT3 (dots) and GooFit (solid) frameworks. Shown are the distributions of the five variables used to parametrize the phase space. Additionally, the normalized pull distributions and p-value are shown. The pulls should follow a normal distribution with a mean of zero (blue line) and a standard deviation of one. The red lines mark the 2σ region.

the subscript $_{12}$ refers to the $\pi^+\ \pi^-$ pair and $_{34}$ to the $\mathrm{K}^+\ \pi^-$ pair. As shown in figure 3, there are no significant differences observed and the pull distribution as well as the p-value indicate that both samples are drawn from the same distribution.

4. Statistical sensitivity to the charm mixing parameters x and y

The novel functionality of GooFit has successfully been used to determine the statistical sensitivity on the charm mixing parameters x and y in a time-dependent amplitude analysis of the decay $\mathrm{D}^0 \to \mathrm{K}^+\pi^-\pi^+\pi^-$. This study did not account for resolution effects, background in the data, and did not allow the model to float. Therefore, the real sensitivity will be worse than shown in table 1. Nevertheless, this study proves the capabilities of the newly implemented extension in GooFit to be fully functional.

Table 1. Summary of the obtained statistical sensitivities of x and y in the case of $x = 0.49\,\%$ and $y = 0.61\,\%$ [11].

Events	Sensitivity of x [%]	Sensitivity of y [%]
20000	0.030	0.031
50000	0.019	0.019
70000	0.016	0.017

5. Performance comparison between CPU and GPU

Lastly, we present a performance comparison of the newly implemented functionality, between the CPU and GPU. Two different test cases are used to study the performance. The first one targets the generation speed of pseudo-events according to a time-dependent amplitude-model. This generation is repeated for three different sample sizes to study the scaling behavior. Secondly, the performance of the fitting procedure is studied, where the scaling behavior is studied by increasing the number of used events in the normalization while leaving the sample size one fits to constant.

These tests are repeated on three different platforms: a server with two Intel Xeon E5-2680 v3 CPUs, each with 12 physical cores that can run two concurrent threads, a NVIDIA K40 GPU and a mid-range mobile gaming GPU NVIDIA GeForce GT 525M. The results are obtained by an average over 5 runs, and listed in tables 2 and 3. They show a significant speedup when utilizing the K40 and even the outdated mid-range mobile graphics card was able to perform surprisingly well compared to the other two platforms, but due to insufficient memory it was not able to complete all tests.

Table 2. Pseudo-event generation according to a time-dependent model using a Monte-Carlo accept/reject method.

	$2\times$ Intel Xeon		NVIDIA	
Events	E5-2680 v3 2.50GHz		GT 525M	K40
	24 Cores	48 Cores	96 Cores	2880 Cores
20 000	179.7 s	156.7 s	195.8 s	25.0 s
50 000	451.9 s	378.7 s	484.6 s	58.8 s
70 000	598.0 s	524.0 s	677.4 s	79.0 s

Table 3. Fit to 100 000 generated pseudo-events, with varying number of points used to calculate the normalization. Fixed model, floating x and y.

	2 × Intel Xeon		NVIDIA	
Points	E5-2680 v3 2.50GHz		GT 525M	K40
	24 Core	48 Cores	96 Cores	2880 Cores
750 000	4.2 s	3.3 s	8.2 s	0.6 s
1 500 000	7.7 s	6.5 s	-	1.0 s
3 000 000	14.8 s	12.4 s	-	1.7 s
6 000 000	30.0 s	22.5 s	-	3.2 s

While the non-linear scaling from 24 to 48 cores was expected as one only increases the logical number of cores by running two threads per core, the expected performance gain from the K40 compared to the GT 525M was less than a priori expected. Using the available NVIDIA profiler, we are able to determine that the source of the throttled performance on the K40 is due to memory latency. We hope to reduce this in the future by reducing the used memory as well adapting the current memory layout to make memory transfers more efficient.

6. Summary

In conclusion, we have presented a novel extension to the GooFit framework which allows for performing a time-dependent amplitude analysis of a pseudoscalar meson decaying into four pseudo-scalar final states. Additionally, this extension allows the user to generate pseudo-events according to a previously defined time-dependent amplitude model. This functionality was successfully validated by comparing the results to an existing software package and furthermore used to study the sensitivity to the charm mixing parameters in the decay $D^0 \to K^+\pi^-\pi^+\pi^-$. Lastly, it is shown that there is a significant speedup gained by utilizing the GPU, while an even bigger performance gain is forseen once the memory layout in GooFit has been adapted to minimize memory latency on high performance GPUs like the K40.

The GooFit package can be found on GitHub at `https://github.com/GooFit`

Acknowledgments

I would like to thank the authors and maintainers of the MINT and MINT3 framework, P. d'Argent, T.D. Evans and J. Rademacker, as their work and support has been most helpful in implementing the presented extension to GooFit.

Work sponsored by the Wolfgang Gentner Programme of the Federal Ministry of Education and Research.

The development of this extension has been in part supported by the National Science Foundation under grant number PHY-1414736.

NVidia provided K40 GPUs for our use through its University Partnership program.

References

[1] Andreassen R, Meadows B T, de Silva M, Sokoloff M D and Tomko K 2014 *Journal of Physics: Conference Series* **513** 052003

[2] Andreassen R E, de Silva W M, Meadows B T, Sokoloff M D and Tomko K A 2014 *IEEE Access* **2** 160–176

[3] The GooFit package `https://github.com/GooFit/GooFit`

[4] The Thrust library `https://thrust.github.io/`

[5] Verkerke W and Kirkby D P 2003 *eConf* **C0303241** (*Preprint* physics/0306116)

[6] Sozzi M S 2008 *Discrete symmetries and CP violation: From experiment to theory*

[7] The MCBooster library https://github.com/MultithreadCorner/MCBooster

[8] Alves Jr A A *et al.* 2016 MCBooster: a library for fast Monte Carlo generation of phase-space decays in massively parallel platforms. http://indico.cern.ch/event/505613/contributions/2230884/

[9] LHCb Collaboration Analysis in progress

[10] Artuso M *et al.* (CLEO) 2012 *Phys. Rev.* **D85** 122002 (*Preprint* 1201.5716)

[11] 2016 Global fit for D^0-$\overline{D}^0$ Mixing https://www.slac.stanford.edu/xorg/hfag/charm/CHARM15/results_mix_cpv.html