

TopoAna: A generic tool for the topology analysis of inclusive Monte-Carlo samples in high energy physics experiments

Xingyu Zhou^{a,*}, Shuxian Du^b, Gang Li^c, Chengping Shen^{d,*}

^a*School of Physics, Beihang University, Beijing 100191, China*

^b*School of Physics and Microelectronics, Zhengzhou University, Zhengzhou 450000, China*

^c*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China*

^d*Key Laboratory of Nuclear Physics and Ion-beam Application (MOE) and Institute of Modern Physics, Fudan University, Shanghai 200443, China*

Abstract

Inclusive Monte-Carlo samples are indispensable for signal selection and background suppression in many high energy physics experiments. A clear knowledge of the topology of the samples, including the categories of physics processes and the number of processes in each category, is a great help to investigating signals and backgrounds. To help analysts get the topology information from the raw data of the samples, we develop a topology analysis program, TopoAna, with C++, ROOT, and LaTeX. The program implements the functionalities of component analysis and signal identification by recognizing, categorizing, counting, and tagging events. Independent of specific software frameworks, the program is applicable to many experiments. At present, it has come into use in three e^+e^- colliding experiments: the BESIII, Belle, and Belle II experiments. The use of the program in other experiments is also prospective.

Keywords:

topology analysis; component analysis; signal identification; inclusive Monte-Carlo samples; high energy physics experiments

1. Introduction

One of the most important tasks in the data analysis of high energy physics experiments is to select signals, or in other words, to suppress backgrounds. As for the task, inclusive Monte-Carlo (MC) samples¹ are extremely useful, in that they provide basic, though not perfect, descriptions of the signals and/or backgrounds involved. However, due to the similarities between signals and some backgrounds, it usually takes efforts to establish a set of selection criteria that retain a high signal efficiency and meanwhile keep a low background level. Further optimization of preliminary criteria is often needed in the process. Under the circumstances, a comprehensive understanding of the samples is required. In particular, a clear knowledge of the topology of

*Corresponding author.

E-mail address: zhouxy@buaa.edu.cn, shencp@fudan.edu.cn

¹The samples are referred to as “generic MC samples” in the Belle and Belle II experiments.

the samples is quite helpful. To be specific, the topology information includes the categories of physics processes and the number of processes in each category, involved both in the entire samples and in the individual events. Here, the physics process could be a complete production and decay process involved in an event, or merely a part of it, such as the decay of an intermediate resonance. With the information, one can figure out the main backgrounds (especially the peaking ones), and optimize the selection criteria further by analyzing the differences between the main backgrounds and the signals.

The analysis of the topology information described above is a sort of component analysis. It is complex since it has to classify the physics processes. Another sort of topology analysis often required in practice is signal identification, which is relatively simple because it only aims to search for certain processes of interests. Mostly, signal and background events coexist in the inclusive MC samples. It is meaningful to differentiate them in such cases. The identified signal events can be used to make up a signal sample (removed to avoid repetition) in the absence (presence) of specialized signal MC sample. Occasionally, we have to pick out some decay branches in order to re-weight them according to new theoretical predictions or updated experimental measurements. Signal identification also plays a part in this occasion.

However, since the raw topology data of inclusive MC samples is counter-intuitive, diverse, and overwhelming, it is difficult for analysts to check the topology of the samples directly. To help them obtain the information quickly and easily, a topology analysis program called TopoAna is developed with C++, ROOT [1], and LaTeX. Here, C++ is the programming language, root is the C++ based data analysis software universally used in high energy physics experiments, and LaTeX is used for generating pdf documents containing the topology information. The program implements the functionalities of component analysis and signal identification by recognizing, categorizing, counting, and tagging events accurately, and exports the obtained topology information to plain text, tex source, pdf, and root files clearly.

The program is applicable to inclusive MC samples at any data analysis stage of high energy physics experiments. In the overwhelming majority of situations, it is run over the samples which have undergone some selections, in order to examine the signals and backgrounds in the selected samples as well as the effect of the imposed selections. In such situations, the results of topology analysis are usually used together with other quantities for physics analysis. In spite of this, applying the program to the samples without undergoing any selection facilitates us to validate the generators and decay cards that produce the samples and helps novices get familiar with the topology of the samples.

Not relying on any specific software frameworks, the program applies to many high energy physics experiments. At first, the program was developed for the BESIII experiment, an experiment in the τ -Charm energy region with abundant research topics under study [2, 3]. Then, it was extended substantially for the Belle II experiment, which is primarily dedicated to search for physics beyond the Standard Model in the flavor sector and has already started data taking in the recent two years [4]. Besides, the program has also been tried and used in the Belle experiment, the predecessor of the Belle II experiment, where some physics studies are still ongoing [5].

This paper gives an essential description of TopoAna. Its structure is organized as follows: Section 2 introduces the basics of the program; Sections 3 and 4 expatiate the two sorts of functionalities of the program — component analysis and signal identification, respectively; Sections 5 and 6 present some common settings and auxiliary facilities for the executing of the program, respectively; Section 7 summarizes the paper.

67 2. Basics of the program

68 This section introduces the basics of the program, including the package, input, execu-
69 tion, and output of the program. The package implements the program via a C++ class called
70 “topoana” and a main function invoking the class. Compiling the package creates the executable
71 file of the program, that is, “topoana.exe”. To execute the program, we have to first obtain the
72 input data of the program, namely the raw topology data of the inclusive MC samples, with some
73 interfaces to the program in the software systems of the corresponding experiments. Normally,
74 the input data contain all the topology information of the samples. With the data, all sorts of the
75 topology analysis presented in the paper can be performed.

76 To carry out the topology analysis desired in our work, we have to provide some necessary
77 input, functionality, and output information to the program. The information is required to be
78 filled in the setting items designed and implemented in the program, and the items have to be put
79 in a plain text file named with a suffix “.card”. With the card file, one can execute the program
80 with the command line: “topoana.exe cardFileName”, where the argument “cardFileName” is
81 optional and its default value is “topoana.card”. After the execution of the program, we can
82 examine the results of topology analysis in the output files and use them to analyze other experi-
83 mental quantities. The results help us gain a better understanding of the signals and backgrounds
84 and are conducive to carrying our work forward. In the next four subsections, we will present the
85 package, input, execution, and output of the program in detail, with each part in one subsection.

86 2.1. Package of the program

87 The package consists of six directories — “include”, “src”, “bin”, “share”, “examples”, and
88 “utilities” — and three files — “README.md”, “Configure”, and “Makefile”. While the direc-
89 tory “include” only includes one header file “topoana.h”, the directory “src” contains sixty source
90 files “*.cpp” as well as a script file “topoana.C”. Despite the largeness of the program, only one
91 class “topoana” is defined for all of its functionalities. The class is declared in “topoana.h”,
92 implemented in “*.cpp” files, and invoked in “topoana.C”.

93 The file “topoana.card” under the directory “share” saves all the items which are developed to
94 set user specified information for the execution of the program. One can refer to the file when fill-
95 ing in the cards for their own needs. Some plain text files “pid_3pchrg.txtprnm_texpnm_iccp.dat_*”
96 are also included in the directory “share”. They store the basic information of the particles used
97 in the program. The suffixes of their names indicate the experiments they apply to. One should
98 rename the file for his/her experiment to “pid_3pchrg.txtprnm_texpnm_iccp.dat” before he/she
99 employs the program. Besides, the directory “share” also contains three LaTeX style files “
100 geometry.sty”, “ifxetex.sty”, and “makecell.sty”, which are invoked by the program for generat-
101 ing pdf files. The directory “examples” includes plenty of detailed examples, particularly those
102 involved in this paper, and the directory “utilities” contains some useful bash scripts.

103 The file “README.md” briefly introduces how to install and use the program. To set up
104 the program, one should first set the package path with the command “./Configure”. Output of
105 the command are the guidelines for manually adding the absolute path of “topoana.exe” to the
106 environment variable “PATH”, in order to execute it without any path. The second step to set
107 up the program is executing the command “make”. This command compiles the header, source,
108 and script files into the executable file “topoana.exe” under the directory “bin”, according to the
109 rules specified in the “Makefile”. Notably, executing the command once is sufficient, unless the
110 header, source, or script files are updated or revised, or the package is moved.

111 2.2. Input of the program

The input of the program is one or more root (TFile [6]) files including a TTree [7] object which contains raw topology data of the inclusive MC samples under study. To be specific, the data in each entry of the TTree object consists of the following three ingredients associated with the particles produced in an event of the samples: the number of particles, PDG [8] codes of particles, and mother indices of particles. Notably, the particles do not include the initial state particles (e^+ and e^- in e^+e^- colliding experiments), which are default and thus omitted. Besides, the indices of particles are integers starting from zero (included) to the number of particles (excluded); they are obvious and hence not taken as an input ingredient for topology analysis. Equation (1) shows an example of the data.

$$\begin{aligned}
 \text{nMCGen} &= 63 \\
 \text{MCGenPDG} &= 300553, \\
 &\quad -511, 511, -433, 421, 211, 22, -413, 111, 111, 113, \\
 &\quad 211, -431, 22, -323, 213, -421, -211, 22, 22, 22, \\
 &\quad 22, 211, -211, 333, 11, -12, 22, -311, -211, 211, \\
 &\quad 111, 221, 331, 321, -321, 310, 22, 22, 111, 111, \\
 &\quad 111, 111, 111, 221, 111, 111, 22, 22, 22, 22, \\
 &\quad 22, 22, 22, 22, 22, 22, 22, 22, 22, 22, \\
 &\quad 22, 22 \\
 \text{MCGenMothIndex} &= 0, \\
 &\quad 0, 0, 1, 1, 1, 1, 2, 2, 2, 2, \\
 &\quad 2, 3, 3, 4, 4, 7, 7, 8, 8, 9, \\
 &\quad 9, 10, 10, 12, 12, 12, 12, 14, 14, 15, \\
 &\quad 15, 16, 16, 24, 24, 28, 31, 31, 32, 32, \\
 &\quad 32, 33, 33, 33, 36, 36, 39, 39, 40, 40, \\
 &\quad 41, 41, 42, 42, 43, 43, 44, 44, 45, 45, \\
 &\quad 46, 46
 \end{aligned} \tag{1}$$

In the example, nMCGen, MCGenPDG, and MCGenMothIndex are the names of the TBranches [9] used for storing the three ingredients. The complete physics process contained in the data is displayed in Eq. (2).

$$\begin{array}{llll}
 0 & e^+e^- \rightarrow \Upsilon(4S), & -1 & 9 \quad \rho^+ \rightarrow \pi^0\pi^+, & 6 \\
 1 & \Upsilon(4S) \rightarrow B^0\bar{B}^0, & 0 & 10 \quad K^{*-} \rightarrow \pi^-\bar{K}^0, & 6 \\
 2 & B^0 \rightarrow \pi^0\pi^0\rho^0\pi^+D^{*-}, & 1 & 11 \quad D_s^- \rightarrow e^-\bar{\nu}_e\phi\gamma, & 7 \\
 3 & \bar{B}^0 \rightarrow \pi^+D^0D_s^{*-}\gamma, & 1 & 12 \quad \eta \rightarrow \pi^0\pi^0\pi^0, & 8 \\
 4 & \rho^0 \rightarrow \pi^+\pi^-, & 2 & 13 \quad \eta' \rightarrow \pi^0\pi^0\eta, & 8 \\
 5 & D^{*-} \rightarrow \pi^-\bar{D}^0, & 2 & 14 \quad \bar{K}^0 \rightarrow K_S^0, & 10 \\
 6 & D^0 \rightarrow \rho^+K^{*-}, & 3 & 15 \quad \phi \rightarrow K^+K^-, & 11 \\
 7 & D_s^{*-} \rightarrow D_s^-\gamma, & 3 & 16 \quad \eta \rightarrow \gamma\gamma, & 13 \\
 8 & \bar{D}^0 \rightarrow \eta\eta', & 5 & 17 \quad K_S^0 \rightarrow \pi^0\pi^0 & 14
 \end{array} \tag{2}$$

112 Here, the decay branches in the process are placed into two blocks in order to make full use
 113 of the page space. In both blocks, the first, second, and third columns are the indices, textual
 114 expressions, and mother indices of the decay branches. Notably, all the decay branches of $\pi^0 \rightarrow$
 115 $\gamma\gamma$ are omitted in Eq. (2) in order to make the process look more concise. Considering π^0 has a
 116 very large production rate and approximatively 99% of it decays to $\gamma\gamma$, the program is designed
 117 to discard the decay $\pi^0 \rightarrow \gamma\gamma$ by default at the early phase of processing the input data (see
 118 Section 5.1.2 for the setting item to alter the behavior). Since the topology diagram of such a
 119 process looks like a tree, we refer to the complete processes as decay trees. Obviously, the data

do not show the structure automatically. Thus, we need the program to do the topology analysis work.

It is recommended to save the input data in the TTree object together with other quantities for physics analyses, in order to facilitate the examination of the distributions of the quantities with the topology information. The input data can be stored in several types. Normally, the number of particles can be simply stored in a TBranch as a scalar integer, while the PDG codes of particles, as well as the mother indices of particles, can be stored in a TBranch as an array of integers, in a TBranch as a vector of integers, or in a group of TBranches as multiple scalar integers. In the analysis software of the Belle II experiment, double-precision variables are used uniformly to store all the quantities involved in the experiment, and TBranches are not recommended to store arrays and vectors in order to use other tools such as NumPy [10] and pandas [11]. In such a case, we have to store the number of particles in a TBranch as a scalar double-precision number, and store the PDG codes of particles, as well as the mother indices of particles, in a group of TBranches as multiple scalar double-precision numbers. Summing up the above, we have mentioned four storage types of the input information. For the sake of simplification, we refer to them with the following acronyms: AOI, VOI, MSI, and MSD, which are short for array of integers, vector of integers, multiple scalar integers, and multiple scalar double-precision numbers, respectively. All of the storage types are supported by the program, and their acronyms will be used in the related item of the card file (see next subsection for details).

It is easy to get the input of the program within the software framework of high energy physics experiments. In order to facilitate its use, we have developed the interfaces of the program to the software systems of the BESIII, Belle, and Belle II experiments. Similar interfaces for other experiments can also be implemented with ease. Beyond the scope of the paper, we will not discuss the details of the interfaces here.

2.3. Execution of the program

To execute the program, we have to first configure some necessary setting items in a card file, and then run the program with the command line: “topoana.exe cardFileName”. This subsection introduces the essential items for the input, basic functionality, and output of the program. More items that can be set in the card file will be described in the following three sections. Sections 3 and 4 expatiate the available items for the functionalities of the program, and Section 5 presents the optional items for the common settings to control the execution of the program.

An example of the card file containing the essential items is shown as follows.

```
# The following six items set the input of the program.

% Names of input root files
{
  ../input/jpsi_1.root
  ../input/jpsi_2.root
}

% TTree name
{
  evt
}

% Storage type of input raw topology data (Four options: AOI, VOI, MSI, and MSD. Default: AOI)
{
```

```

169     AOI
170 }
171
172 % TBranch name of the number of particles (Default: nMCGen)
173 {
174     Nmcpes
175 }
176
177 % TBranch name of the PDG codes of particles (Default: MCGenPDG)
178 {
179     Pid
180 }
181
182 % TBranch name of the mother indices of particles (Default: MCGenMothIndex)
183 {
184     Midx
185 }
186
187 # The following item sets the basic functionality of the program.
188
189 % Component analysis — decay trees
190 {
191     Y
192 }
193
194 # The following item sets the output of the program.
195
196 % Main name of output files (Default: Main name of the card file)
197 {
198     jpsi.ta
199 }

```

In the card file, “#”, “%”, and the pair of “{” and “}”, are used for commenting, prompting, and grouping, respectively. The first six, seventh, and last one items are set for the input, basic functionality, and output of the program, respectively.

The first item sets the names of the input root files. The names ought to be input one per line without tailing characters, such as comma, semicolon, and period. In the names, both the absolute and relative paths are allowed and wildcards “[]?*” are supported, just like those in the root file names input to the method Add() of the class TChain [12]. The second item specifies the TTree name. The third item tells the program the storage type of the raw input topology data, and the input should be one of the following four acronyms: AOI, VOI, MSI, and MSD, as we introduce in the previous subsection. The following three items set the TBranch names of the three ingredients of the raw input topology data. Of the first six items, the former two are indispensable, whereas the latter four can be removed or left empty if the input values are identical to the default values indicated in their prompts. Besides, the latter four items can be moved to the underlying card file, which is developed for frequently used items and will be introduced in Section 6.1, because the input values are usually fixed for a user or a group of users, though they might be different from the default values.

The seventh item sets the basic functionality of the program, namely the component analysis over decay trees. The item can be replaced or co-exist with other functionality items expatiated in Sections 3 and 4. Here, we note that at least one functionality item has to be specified explicitly in the card file, otherwise the program will terminate soon after its start because no topology analysis to be performed is set up.

222 The last item specifies the main name of the output files. Though in different formats, the files
 223 are denominated with the same main name for the sake of uniformity. They will be introduced
 224 at length in the next subsection. This item is also optional, with the main name of the card file
 225 as its default input value. It is a good practice to first denominate the card file with the desired
 226 main name of the output files and then remove this item or leave it empty.

227 To provide a complete description, we list and explain all the essential items in the paragraphs
 228 above. However, in practical uses, we suggest removing the optional items if the input values are
 229 identical to the default ones, or moving them to the underlying card file if the input values are
 230 fixed for most of your use cases. In this way, the contents of the card file will become much more
 231 concise, making the use of the program easier and quicker. For example, only the following two
 232 items are used to set the essential information in Sections 3, 4, and 5.

```

233
234 % Names of input root files
235 {
236     ../input/mixed_1.root
237     ../input/mixed_2.root
238 }
239
240 % TTree name
241 {
242     evt
243 }
244

```

245 Besides, all the items in the program, also including those to be introduced in the following sec-
 246 tions, are not required to be filled in the card files in a certain order. Nonetheless, we recommend
 247 filling them in a logical order for clearness.

248 During the execution of the program, some standard output and error messages are printed to
 249 the screen to provide some information on the input, progress, and output of the program, as well
 250 as the possible problems and proposed solutions to them. The standard output messages include
 251 the following four parts: (1) the values of the items with active inputs; (2) the total number of
 252 entries contained in the input root files and the progress of the program to process these entries;
 253 (3) the information output by the pdflatex command when it compiles the tex source file to get
 254 the pdf file; (4) and the hints on the output of the program. The standard error messages are
 255 prompted with “ERROR” and “INFOR” in order to differentiate themselves from the standard
 256 output messages. The messages started with “ERROR” point out the problems encountered by
 257 the program directly, while those started with “INFOR” give more information on the problems
 258 as well as some guidelines on the solutions to the problems.

259 2.4. Output of the program

260 The program gains the topology information from input data, and saves the information to
 261 output files. As mentioned in Section 1, the information includes the categories of physics pro-
 262 cesses and the number of processes in each category, involved both in entire samples and in
 263 individual events. We refer to the information at the sample level as topology maps. In the
 264 topology maps, we assign an integer to each category of physics processes as its index. We term
 265 the indices of processes as well as the numbers of processes involved in each category in the
 266 individual events as topology tags.

267 The program outputs topology maps to three different files: one plain text file, one tex source
 268 file, and one pdf file, with the same main name specified in the card file. For instance, the
 269 three files are “jpsi.ta.txt”, “jpsi.ta.tex”, and “jpsi.ta.pdf” in the example. Although in different
 270 formats, the three files have the same information. The pdf file is the easiest to read. It is

converted from the tex source file with the command `pdflatex`. The tex source file is convenient to us if we want to change the style of the pdf file to our taste and when we need to copy and paste (parts of) the topology maps to our slides, papers, and so on. For example, all of the tables displaying topology maps in this paper are taken from associated tex source files. The plain text file has its own advantage, because the topology maps in it can be checked with text processing commands as well as text editors, and can be used on some occasions as input to the functionality items (see Sections 3 and 4 for details) of another card file.

In addition to the three files for topology maps, one or more root files are output to save topology tags. The root files only include one TTree object, which is entirely the same as that in the input root files, except for the topology tags inserted in all of its entries. The number of root files depends on the size of output data. The program switches to one new root file whenever the size of the TTree object in memory exceeds 3 GB. In the case of the size less than 3 GB, only one root file is output. While the sole or first root file has the same main name as the three files above, more possible root files are denominated with the suffix “_n” (n=1, 2, 3, and so on) appended to the main name. In the example, the first root file is “jpsi.ta.root”, and more possible root files would be “jpsi.ta.1.root”, “jpsi.ta.2.root”, “jpsi.ta.3.root”, and so on.

In the example of the previous subsection, the program conducts its basic functionality, namely the component analysis over decay trees. From the 100000 events of the input sample, the program recognizes 17424 decay trees and outputs all of them to the txt, tex, and pdf files. Table 1 shows only the top ten decay trees and their respective final states in the obtained topology map. In the table, “rowNo”, “iDcyTr”, “nEtr”, and “nCEtr” are abbreviations for the row number, index of decay tree, number of entries of decay tree, and number of the cumulative entries from the first to the current decay trees, respectively. The values of “iDcyTr” are assigned from small to large in the program but listed according to the values of “nEtr” from large to small in the table. This is the reason why they are not in natural order like the values of “rowNo”. In Section 2.2, we mention that the decay $\pi^0 \rightarrow \gamma\gamma$ is ignored by the program at the early phase of processing the input data. As a result, it does not show itself in the table.

Table 1: Top ten decay trees and their respective final states.

rowNo	decay tree	decay final state	iDcyTr	nEtr	nCEtr
1	$J/\psi \rightarrow \mu^+ \mu^-$	$\mu^+ \mu^-$	6	5269	5269
2	$J/\psi \rightarrow e^+ e^-$	$e^+ e^-$	4	4513	9782
3	$J/\psi \rightarrow \pi^0 \pi^+ \pi^- \pi^-$	$\pi^0 \pi^+ \pi^- \pi^-$	0	2850	12632
4	$J/\psi \rightarrow \pi^0 \pi^+ \pi^+ \pi^- \pi^-$	$\pi^0 \pi^+ \pi^+ \pi^- \pi^-$	2	1895	14527
5	$J/\psi \rightarrow \pi^0 \pi^+ \pi^- K^+ K^-$	$\pi^0 \pi^+ \pi^- K^+ K^-$	20	1698	16225
6	$J/\psi \rightarrow \rho^+ \rho^- \omega, \rho^+ \rightarrow \pi^0 \pi^+, \rho^- \rightarrow \pi^0 \pi^-, \omega \rightarrow \pi^0 \pi^+ \pi^-$	$\pi^0 \pi^0 \pi^0 \pi^+ \pi^- \pi^-$	19	1453	17678
7	$J/\psi \rightarrow e^+ e^- \gamma^f$	$e^+ e^- \gamma^f$	70	1222	18900
8	$J/\psi \rightarrow \pi^0 \pi^0 \pi^+ \pi^- \pi^-$	$\pi^0 \pi^0 \pi^+ \pi^- \pi^-$	127	1161	20061
9	$J/\psi \rightarrow \pi^0 \pi^+ \pi^+ \pi^+ \pi^- \pi^- \pi^-$	$\pi^0 \pi^+ \pi^+ \pi^+ \pi^- \pi^- \pi^-$	234	836	20897
10	$J/\psi \rightarrow \pi^0 \pi^0 \pi^+ \pi^- \gamma^F$	$\pi^0 \pi^0 \pi^+ \pi^- \gamma^F$	43	792	21689

In the table, “iDcyTr” is the topology tag for decay trees. Thus, it is also saved in the TTree objects of the output root file, together with other quantities for physics analysis. Therefore, it can be used to pick out the entries of specific decay trees and then examine the distributions of the other quantities over the decay trees. In the example, besides the raw topology data, only

a random variable following standardized normal distribution, namely X , is stored in the input root files and thus copied by default to the output root file. Though not a genuine variable for physics analysis, X is quite good to illustrate the usage of the topology tag. Figure 1 shows the distribution of X accumulated over the top ten decay trees. The figure is drawn with the root script

```
examples/in_the_paper/ex_for_tb.01/draw_X/v2/draw_X.C,
```

where, for example, a statement equivalent to

```
chain->Draw("X >>h0", "iDcyTr==6")
```

is used to import X over the decay tree $J/\psi \rightarrow \mu^+ \mu^-$ from the output root file to the histogram named h0. With such a figure, we can clearly see the contribution of each decay tree. Particularly, we can get to know whether a decay tree has a peak contribution or a contribution mainly distributed in a different region. Based on these distributions, we can get a better understanding of our signals and backgrounds, and thus optimize event selection criteria by applying new requirements on the displayed quantities.

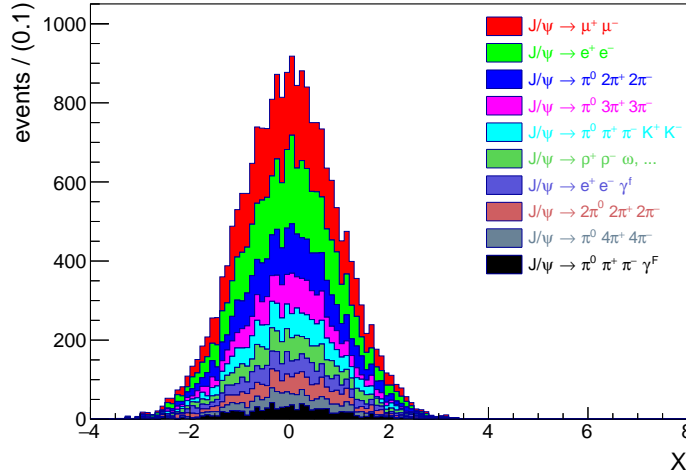


Figure 1: Distribution of X accumulated over the top ten decay trees. In the legend entry “ $J/\psi \rightarrow \rho^+ \rho^- \omega, \dots$ ”, the dots “ \dots ” represent the secondary decay branches: $\rho^+ \rightarrow \pi^0 \pi^+$, $\rho^- \rightarrow \pi^0 \pi^-$, $\omega \rightarrow \pi^0 \pi^+ \pi^-$.

3. Component analysis

Component analysis is the primary functionality of the program. It is performed over decay trees in the previous example. In addition, it can be carried out as follows: over decay initial-final states; with specified particles to check their decay branches, production branches, mothers, cascade decay branches, and decay final states; with specified inclusive decay branches to examine their exclusive components; and with specified intermediate-resonance-allowed (IRA) decay branches to investigate their inner structures. This section introduces the nine (five for specified

particles) kinds of component analysis, with each in a subsection. For each kind of component analysis, one item is designed and implemented in the program to set related parameters. In each subsection, we take an example to demonstrate the corresponding setting item and show the resulting topology map. For easy exposition, all of the essential topology tags involved in the component analysis functionalities are presented in another separate subsection, namely the last subsection.

Similar to the case over decay trees, to perform the component analysis over decay initial-final states, we only need to input an positive option “Y” to the corresponding item. Different from the former two kinds, to carry out the latter seven kinds of component analysis, we have to explicitly specify one or more desired particles, inclusive decay branches, or IRA decay branches in the associated items. In the following examples, two particles or decay branches are set to illustrate the use of these items, but only the topology map related to one of them is shown to save space in the paper.

In addition to the indispensable parameters, two sorts of common optional parameters can be set in the items. The first sort is designed for all the nine kinds of component analysis to restrict the maximum number of components output to the txt, tex and pdf files. Without the optional parameters, all components will be output. This is fine if the number of components is not massive. In cases of too many (around ten thousand or more) components, it takes a long time for the program to output the components to the txt and tex files as well as to get the pdf file from the tex file. In such cases, it also takes up a large disk space to save these components in the output files. Considering further that the posterior components are generally unimportant and our time and energy to examine them are limited, it is better to set a maximum to the number of output components. To save space in the paper, we set the maximum number to five in the following examples.

The second sort of optional parameters are developed for the latter seven kinds of component analysis to assign meaningful aliases to the specified particles, inclusive decay branches, and IRA decay branches. By default, the indices 0, 1, 2, and so on are used to tag the particles and decay branches in the names of the TBranches appended in the TTree object of the output root files. This is fine, but it is significative to replace the indices with meaningful aliases, particularly in cases of many specified particles or decay branches.

3.1. Decay trees

Component analysis over decay trees is the basic kind of component analysis that has already been widely performed in the BESIII experiment. The following example shows the associated item with the maximum number of output components set to five. In the item, a third parameter is also filled and set to “Y”. With the setting, the decay final states in the output pdf file are put under their respective decay trees, rather than in a column next to that for decay trees. It is recommended to use this optional parameter in cases there are too many (about ten or more) particles in some final states. Here, we note that the symbol “—” can be used as a placeholder for the maximum number of output components, if only the third parameter is desired.

```
% Component analysis — decay trees
{
  Y  5  Y
}
```

Table 2 shows the decay trees. In the table, while the first five decay trees are listed exclusively

in the main part, the rest decay trees are only summarized inclusively at the bottom row. In the symbolic expressions of decay initial-final states, the dashed right arrow ($--\rightarrow$) instead of the plain right arrow (\rightarrow) is used, in order to reflect that the initial state does not necessarily decay to the final states in a direct way. Similarly, it is also used in the symbolic expressions of IRA decay branches, which will be introduced in Section 3.9.

Table 2: Decay trees and their respective initial-final states.

rowNo	decay tree (decay initial-final states)	iDcyTr	nEtr	nCEtr
1	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow e^+ \nu_e D^{*-} \gamma^F, \bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}, D^{*-} \rightarrow \pi^- \bar{D}^0,$ $D^{*+} \rightarrow \pi^+ D^0, \bar{D}^0 \rightarrow \pi^0 \pi^- K^+, D^0 \rightarrow \pi^0 \pi^+ K^-$ $(\Upsilon(4S) --\rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu \pi^0 \pi^+ \pi^- \pi^- K^+ K^- \gamma^F)$	20870	3	3
2	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow \pi^0 \pi^+ \pi^- \rho^- D^-, \bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}, \rho^- \rightarrow \pi^0 \pi^-,$ $D^- \rightarrow \pi^- \pi^- K^+, D^{*+} \rightarrow \pi^+ D^0, D^0 \rightarrow K_L^0 \pi^+ \pi^-$ $(\Upsilon(4S) --\rightarrow \mu^- \bar{\nu}_\mu \pi^0 \pi^0 K_L^0 \pi^+ \pi^+ \pi^- \pi^- \pi^- \pi^- K^+)$	5295	2	5
3	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow \mu^+ \nu_\mu D^{*-}, \bar{B}^0 \rightarrow e^- \bar{\nu}_e D^+, D^{*-} \rightarrow \pi^- \bar{D}^0,$ $D^+ \rightarrow e^+ \nu_e \bar{K}^*, \bar{D}^0 \rightarrow \pi^0 \pi^+ \pi^- K_S^0, \bar{K}^* \rightarrow \pi^0 \bar{K}^0, K_S^0 \rightarrow \pi^+ \pi^-, \bar{K}^0 \rightarrow K_L^0$ $(\Upsilon(4S) --\rightarrow e^+ e^- \nu_e \bar{\nu}_e \mu^+ \nu_\mu \pi^0 \pi^0 K_L^0 \pi^+ \pi^+ \pi^- \pi^- \pi^-)$	11954	2	7
4	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow e^+ \nu_e D^{*-}, \bar{B}^0 \rightarrow \pi^0 \pi^- \omega D^+, D^{*-} \rightarrow \pi^- \bar{D}^0,$ $\omega \rightarrow \pi^0 \pi^+ \pi^-, D^+ \rightarrow e^+ \nu_e \pi^+ K^-, \bar{D}^0 \rightarrow \pi^0 \pi^- K^+$ $(\Upsilon(4S) --\rightarrow e^+ e^+ \nu_e \nu_e \pi^0 \pi^0 \pi^+ \pi^- \pi^- \pi^- \pi^- K^+ K^-)$	14345	2	9
5	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow \mu^+ \nu_\mu D^{*-}, \bar{B}^0 \rightarrow e^- \bar{\nu}_e D^{*+} \gamma^F, D^{*-} \rightarrow \pi^- \bar{D}^0,$ $D^{*+} \rightarrow \pi^0 D^+, \bar{D}^0 \rightarrow \pi^- K^+, D^+ \rightarrow e^+ \nu_e \bar{K}^*, \bar{K}^* \rightarrow \pi^+ K^-$ $(\Upsilon(4S) --\rightarrow e^+ e^- \nu_e \bar{\nu}_e \mu^+ \nu_\mu \pi^0 \pi^+ \pi^- \pi^- K^+ K^- \gamma^F)$	15332	2	11
rest	$\Upsilon(4S) \rightarrow \text{others (99980 in total)}$ $(\Upsilon(4S) --\rightarrow \text{corresponding to others})$	—	99989	100000

3.2. Decay initial-final states

On some occasions, we need to investigate decay initial-final states. Below is an example demonstrating the related item with the maximum number of output components set to five.

```
% Component analysis — decay initial-final states
{
  Y  5
}
```

The decay initial-final states are displayed in Table 3. The layout of the table is similar to that of Table 2, which shows the decay trees.

Table 3: Decay initial-final states.

rowNo	decay initial-final states	iDcyIFSts	nEtr	nCEtr
1	$\Upsilon(4S) --\rightarrow \mu^+ \nu_\mu \pi^0 \pi^0 \pi^0 \pi^+ \pi^+ \pi^- \pi^- \pi^- \pi^- K^+ K^-$	41	18	18
2	$\Upsilon(4S) --\rightarrow \pi^0 \pi^0 \pi^0 \pi^0 \pi^0 \pi^+ \pi^+ \pi^+ \pi^- \pi^- \pi^- \pi^- K^+ K^-$	887	18	36
3	$\Upsilon(4S) --\rightarrow \mu^- \bar{\nu}_\mu \pi^0 \pi^0 \pi^0 \pi^+ \pi^+ \pi^+ \pi^- \pi^- \pi^- \pi^- K^+ K^-$	3350	18	54
4	$\Upsilon(4S) --\rightarrow \pi^0 \pi^0 \pi^0 \pi^0 \pi^0 \pi^+ \pi^+ \pi^+ \pi^- \pi^- \pi^- \pi^- K^+ K^-$	1215	17	71
5	$\Upsilon(4S) --\rightarrow \pi^0 \pi^0 \pi^0 \pi^0 \pi^0 K_L^0 \pi^+ \pi^+ \pi^+ \pi^- \pi^- \pi^- \pi^- K^-$	1207	17	88
rest	$\Upsilon(4S) --\rightarrow \text{others (78208 in total)}$	—	99912	100000

3.3. Decay branches of particles

Sometimes, we have to examine the decay branches of certain particles. The following example shows the associated item with the two particles D^{*+} and J/ψ set as research objects. In the item, each row holds the information of a specified particle, and the first, second and third columns are the textual expressions, aliases, and maximum numbers of output components, respectively. As we introduce at the beginning part of this section, the aliases and maximum numbers of output components are both optional. Here, we note that the symbol “—” can be used as a placeholder for an unassigned alias, if only the maximum number of output components is desired.

```
% Component analysis — decay branches of particles
{
  D*+   Dsp   5
  J/psi  Jpsi  5
}
```

Table 4 shows the decay branches of D^{*+} . From the table, only four decay branches of D^{*+} are found in the input inclusive MC sample. Since there is likely one or more cases of D^{*+} decays in one input entry, “nCase” and “nCCase”, instead of “nEtr” and “nCEtr”, are used in the table in order to accurately indicate what we are counting are the numbers of D^{*+} decays, rather than the numbers of entries involving the D^{*+} decays.

Table 4: Decay branches of D^{*+} .

rowNo	decay branch of D^{*+}	iDcyBrP	nCase	nCCase
1	$D^{*+} \rightarrow \pi^+ D^0$	0	31180	31180
2	$D^{*+} \rightarrow \pi^0 D^+$	1	13978	45158
3	$D^{*+} \rightarrow D^+ \gamma$	2	700	45858
4	$D^{*+} \rightarrow \pi^+ D^0 \gamma^F$	3	28	45886

3.4. Production branches of particles

In some cases, we have interest in the production branches of certain particles. Below is an example demonstrating the related item also by taking the two particles D^{*+} and J/ψ as objects of study. The input to this item is the same as that to the above item.

```
% Component analysis — production branches of particles
{
  D*+   Dsp   5
  J/psi  Jpsi  5
}
```

Table 5: Production branches of D^{*+} .

rowNo	production branch of D^{*+}	iProdBrP	nCase	nCCase
1	$\bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}$	9	4154	4154
2	$\bar{B}^0 \rightarrow e^- \bar{\nu}_e D^{*+}$	7	2886	7040
3	$\bar{B}^0 \rightarrow D^{*+} D_s^{*-}$	4	1691	8731
4	$\bar{B}^0 \rightarrow e^- \bar{\nu}_e D^{*+} \gamma^F$	10	1623	10354
5	$\bar{B}^0 \rightarrow \pi^0 \pi^+ \pi^- D^{*+}$	40	1429	11783
rest	others (3272 in total)	—	34103	45886

The production branches of D^{*+} are displayed in Table 5. In the production branches, D^{*+} is marked in blue so as to make it noticeable. From the table, the number of production branches of D^{*+} found in the input sample is 3277, much bigger than 4, which is the number of its decay branches.

3.5. Mothers of particles

Occasionally, we may want to check the mothers of certain particles. The following example shows the associated item also with the two particles D^{*+} and J/ψ set as research objects. The input to this item is identical to those to the two items above.

```
% Component analysis — mothers of particles
{
  D*+  Dsp  5
  J/psi Jpsi 5
}
```

Table 6 shows the mothers of D^{*+} . Notably, the PDG codes of the mother particles, instead of additional indices, are listed in the table, since they are sufficient to tag the mother particles. From the table, six sources of D^{*+} are found in the input sample and the dominant one is the \bar{B}^0 decay.

Table 6: Mothers of D^{*+} .

rowNo	mother of D^{*+}	PDGMoth	nCase	nCCase
1	\bar{B}^0	-511	41751	41751
2	B^0	511	2983	44734
3	D_1^{*+}	20413	455	45189
4	D_1^+	10413	368	45557
5	D_2^{*+}	415	247	45804
rest	others (1 in total)	—	82	45886

3.6. Cascade decay branches of particles

On some occasions, we need to investigate the cascade decay branches of certain particles. Below is an example demonstrating the related item by taking the two particles B^0 and D^0 as objects of study. While the first three columns of the input to this item have the same meanings as those to the three items above, the additional fourth column sets the maximum hierarchy of decay branches to be examined. Here, the hierarchy reflects the rank of a decay branch in a cascade decay branch of one specific particle. For instance, in the following cascade decay branch of B^0 : $B^0 \rightarrow \pi^0 \pi^0 \rho^0 \pi^+ D^{*-}$, $\rho^0 \rightarrow \pi^+ \pi^-$, $D^{*-} \rightarrow \pi^- \bar{D}^0$, $\bar{D}^0 \rightarrow \eta \eta'$, $\eta \rightarrow \pi^0 \pi^0 \pi^0$, $\eta' \rightarrow \pi^0 \pi^0 \eta$, $\eta \rightarrow \gamma \gamma$, the hierarchies of the seven individual decay branches are 1, 2, 2, 3, 4, 4, and 5, respectively. In the example item, the maximum hierarchy of decay branches is set to two for both B^0 and D^0 , and hence only the first two hierarchies of branches in their cascade decays will be investigated. Without such settings, all the branches in their cascade decays will be examined.

```
% Component analysis — cascade decay branches of particles
{
  B0  B0  5  2
  D0  D0  5  2
}
```

The cascade decay branches of B^0 are displayed in Table 7.

Table 7: Cascade decay branches of B^0 (only the first two hierarchies are involved).

rowNo	cascade decay branch of B^0	iCascDcyBrSP	nCase	nCCase
1	$B^0 \rightarrow \mu^+ \nu_\mu D^{*-}, D^{*-} \rightarrow \pi^- \bar{D}^0$	12	2912	2912
2	$B^0 \rightarrow e^+ \nu_e D^{*-}, D^{*-} \rightarrow \pi^- \bar{D}^0$	6	1991	4903
3	$B^0 \rightarrow \mu^+ \nu_\mu D^{*-}, D^{*-} \rightarrow \pi^0 D^-$	70	1283	6186
4	$B^0 \rightarrow e^+ \nu_e D^{*-} \gamma^F, D^{*-} \rightarrow \pi^- \bar{D}^0$	18	1132	7318
5	$B^0 \rightarrow D^{*-} D_s^{*+}, D^{*-} \rightarrow \pi^- \bar{D}^0, D_s^{*+} \rightarrow D_s^+ \gamma$	20	1119	8437
rest	$B^0 \rightarrow \text{others (42074 in total)}$	—	91594	100031

3.7. Decay final states of particles

Sometimes, we have to examine the decay final states of certain particles. The following example shows the associated item also with the two particles B^0 and D^0 set as research objects. The format of the input to the item is the same as that to the above item, but the fourth parameters here are designed to restrict the numbers of final state particles. Without the fourth parameters, all the decay final states of the specified particles will be investigated. In the example, the parameters are set to three for both B^0 and D^0 , and thus only the three-body decay final states of them will be examined.

```
% Component analysis — decay final states of particles
{
  B0  B0  5  3
  D0  D0  5  3
}
```

Table 8 shows the three-body decay final states of D^0 . In the table, π^0 only decays to $\gamma\gamma$; otherwise, it will be replaced with its decay products, resulting in different decay final states of D^0 .

Table 8: Decay final states of D^0 (only three-body final states are involved).

rowNo	decay final state of D^0	iDcyFStP	nCase	nCCase
1	$D^0 \rightarrow \pi^0 \pi^+ K^-$	2	6258	6258
2	$D^0 \rightarrow \mu^+ \nu_\mu K^-$	5	1487	7745
3	$D^0 \rightarrow \pi^0 \pi^+ \pi^-$	1	1162	8907
4	$D^0 \rightarrow K_L^0 \pi^+ \pi^-$	3	1158	10065
5	$D^0 \rightarrow e^+ \nu_e K^-$	11	1148	11213
rest	$D^0 \rightarrow \text{others (24 in total)}$	—	2407	13620

3.8. Inclusive decay branches

In some cases, we are interested in the exclusive components of certain inclusive decay branches. Below is an example demonstrating the related item by taking the two inclusive decay branches $\bar{B}^0 \rightarrow D^{*+} + \text{anything}$ and $B^0 \rightarrow K_S^0 + \text{anything}$ as objects of study. In the item, each row holds the information of a inclusive decay branch, and the first, second, and third columns separated with the symbol “&” are the textual expressions, aliases, and maximum numbers of output components, respectively. As we introduce at the beginning part of this section, the aliases and maximum numbers of output components are both optional. Here, we note that the symbol “_” can be used as a placeholder for an unassigned alias, if only the maximum number of output components is desired.

```
% Component analysis — inclusive decay branches
```

```

491 {
492   B0 --> D*+ & B2Dsp & 5
493   B0 --> K_S0 & B2Ks & 5
494 }

```

495

496 The exclusive components of $B^0 \rightarrow K_S^0 + \text{anything}$ are displayed in Table 9. From the table,
 497 ten exclusive components of the inclusive decay branch are found in the input sample, and the
 498 particles denoted with *anything* are mainly the traditional charmonium states.

Table 9: Exclusive components of $B^0 \rightarrow K_S^0 + \text{anything}$.

rowNo	exclusive component of $B^0 \rightarrow K_S^0 + \text{anything}$	iDcyBrIncDcyBr	nCase	nCCase
1	$B^0 \rightarrow K_S^0 J/\psi$	0	45	45
2	$B^0 \rightarrow K_S^0 \eta_c$	1	40	85
3	$B^0 \rightarrow K_S^0 \psi'$	3	33	118
4	$B^0 \rightarrow K_S^0 \chi_{c1}$	2	20	138
5	$B^0 \rightarrow K_S^0 \chi_{c0}$	4	6	144
rest	$B^0 \rightarrow K_S^0 + \text{others (5 in total)}$	—	9	153

499 3.9. Intermediate-resonance-allowed decay branches

500 Occasionally, we may want to investigate the inner structures of certain IRA decay branches.
 501 The following example shows the associated item with the two IRA decay branches $D^{*+} \rightarrow$
 502 $\pi^0 \pi^+ \pi^+ K^-$ and $J/\psi \rightarrow \pi^0 \pi^+ \pi^-$ set as research objects. Since IRA decay branches look like
 503 inclusive decay branches, the format of the input to the item for IRA decay branches is identical
 504 to that for inclusive decay branches, which is introduced in the previous subsection.

```

505 % Component analysis — intermediate-resonance-allowed decay branches
506 {
507   D*+ --> K- pi+ pi+ pi0 & Dsp2K3Pi & 5
508   J/psi --> pi+ pi- pi0 & Jpsi23Pi & 5
509 }
510
511

```

512 Table 10 shows the exclusive components of $D^{*+} \rightarrow \pi^0 \pi^+ \pi^+ K^-$. From the table, two interme-
 513 diate particles D^0 and D^+ are found in the IRA decay branch, and they decay to $\pi^0 \pi^+ K^-$ and
 514 $\pi^+ \pi^+ K^-$, respectively.

Table 10: Exclusive components of $D^{*+} \rightarrow \pi^0 \pi^+ \pi^+ K^-$.

rowNo	exclusive component of $D^{*+} \rightarrow \pi^0 \pi^+ \pi^+ K^-$	iDcyBrIRADcyBr	nCase	nCCase
1	$D^{*+} \rightarrow \pi^+ D^0, D^0 \rightarrow \pi^0 \pi^+ K^-$	0	3869	3869
2	$D^{*+} \rightarrow \pi^0 D^+, D^+ \rightarrow \pi^+ \pi^+ K^-$	1	1102	4971

515 3.10. Essential topology tags

516 Table 11 lists and interprets all of the essential topology tags involved in the component
 517 analysis functionalities. The topology tag for the component analysis over decay initial-final
 518 states is iDcyIFSts. It has the similar interpretation as iDcyTr and is shown in the third column
 519 of Table 3. For the latter seven kinds of component analysis, there are two sorts of topology tags.
 520 The first sort, such as nPDcyBr_i, records the number of the specified particles or decay branches
 521 found in each entry. The second sort, for example iDcyBrP_i_j, keeps the corresponding index of

Table 11: Essential topology tags involved in each kind of component analysis.

Component analysis kind	Topology tag	Interpretation
Decay trees	iDcyTr	index of decay tree
Decay initial-final states	iDcyIFSts	index of decay initial-final states
Decay branches of particles	nPDcyBr.i	number of particle _i s (or its decay branches)
	iDcyBrP.i.j	index of decay branch of the j th particle _i
Production branches of particles	nPProdBr.i	number of particle _i s (or its production branches)
	iProdBrP.i.j	index of production branch of the j th particle _i
Mothers of particles	nPMoth.i	number of particle _i s (or its mothers)
	PDGMothP.i.j	PDG code of mother of the j th particle _i
Cascade decay branches of particles	nPCascDcyBr.i	number of particle _i s (or its cascade decay branches)
	iCascDcyBrP.i.j	index of cascade decay branch of the j th particle _i
Decay final states of particles	nPDcyFSt.i	number of particle _i s (or its decay final states)
	iDcyFStP.i.j	index of decay final state of the j th particle _i
Inclusive decay branches	nIncDcyBr.i	number of inclusive decay branch _i es
	iDcyBrIncDcyBr.i.j	index of decay branch of the j th inclusive decay branch _i
IRA decay branches	nIRADcyBr.i	number of IRA decay branch _i es
	iDcyBrIRADcyBr.i.j	index of decay branch of the j th IRA decay branch _i

each instance of one specified particle or decay branch found in each entry. The indices are also listed in the third columns of Tables 4 – 10.

In the topology tags, “i” in “_i” is the default index of the specified particle or decay branch, and it ranges from 0 (included) to the number of specified particles or decay branches (excluded). If the alias of the particle or decay branch is also specified, the index “i” will be replaced with the alias. For example, in the component analysis over decay branches of D^{*+} and J/ψ , since “Dsp” and “Jpsi” are set as their aliases, the specialised topology tags nPDcyBr_Dsp and nPDcyBr_Jpsi, instead of the default ones nPDcyBr_0 and nPDcyBr_1, are used to store the numbers of D^{*+} and J/ψ (or their decay branches) found in each entry.

In addition, “j” in “_j” is the default index of the particle or decay branch found in an entry, and it ranges from 0 (included) to the sample-level maximum of the number of the particles or decay branches found in each entry (excluded). For example, in the component analysis over decay branches of D^{*+} , the sample-level maximum of the number of D^{*+} (or its decay branches) found in each entry is two, and thus the indices of the D^{*+} decay branches are stored in the topology tags iDcyBrP_Dsp_0 and iDcyBrP_Dsp_1. The indices range from 0 (include) to the number of the categories of the D^{*+} decay branches found in the samples. In the entries with only one D^{*+} , iDcyBrP_Dsp_1 is assigned with the default value -1; in the entries which have no D^{*+} , the default value -1 is assigned to both iDcyBrP_Dsp_0 and iDcyBrP_Dsp_1. We note that different from all other indices, PDGMoth.i.j has the default value 0, instead of -1.

4. Signal identification

Signal identification is the other functionality of the program. Though it is a relatively simple functionality, it can help us identify the signals we desire directly, quickly, and easily. At present, the following eight kinds of signals can be identified with the program: (1) decay trees, (2) decay initial-final states, (3) particles, (4) (regular) decay branches, (5) cascade decay branches, (6) inclusive decay branches, (7) inclusive cascade decay branches, and (8) IRA decay branches. For each kind of signals, one item is developed to specify related parameters. This section introduces

the eight kinds of signal identification, with each in a subsection. In each subsection, we take an example to demonstrate the related setting item and show the obtained topology map. For easy exposition, all of the essential topology tags involved in the signal identification functionalities are presented in another separate subsection, that is, the last subsection.

Similar to the cases of the latter seven kinds of component analysis, one or more signals can be specified in each of the signal identification items, and two signals are set in the following examples to illustrate the use of the items. Besides, meaning aliases can also be optionally assigned to the specified signals so as to better tag them in the names of the TBranches appended in the TTree object of the output root files.

4.1. Decay trees

Sometimes, we need to identify certain decay trees. The following example shows the associated item with the first two decay trees listed in Table 2 set as signals. In the item, each row holds a decay branch in the decay trees, and the first, second, and third columns separated with the symbol “&” are the indices, textual expressions, and mother indices of the decay branches, respectively. In addition, the decay branches with index 0 indicate the beginning of new decay trees, and their mother indices are equal to -1, suggesting they have no mother branches because they are the first decay branches of the decay trees. Besides, a name of each decay tree can be optionally filled in the fourth column of its first decay branch. Similar to the third parameter in the item for the component analysis over decay trees (see Section 3.1), a “Y” can be optionally filled in the fifth column of the first decay branch of the first decay tree, to adjust the positions of decay final states in the output pdf file.

```
% Signal identification — decay trees
{
  0 & Upsilon(4S) --> B0 anti-B0 & -1 & 1stDcyTrInTb2 & Y
  1 & B0 --> e+ nu_e D*- gamma & 0
  2 & anti-B0 --> mu- anti-nu_mu D*+ & 0
  3 & D*- --> pi- anti-D0 & 1
  4 & D*+ --> pi+ D0 & 2
  5 & anti-D0 --> pi0 pi- K+ & 3
  6 & D0 --> pi0 pi+ K- & 4
  0 & Upsilon(4S) --> B0 anti-B0 & -1 & 2ndDcyTrInTb2
  1 & B0 --> pi0 pi+ pi- rho- D- & 0
  2 & anti-B0 --> mu- anti-nu_mu D*+ & 0
  3 & rho- --> pi0 pi- & 1
  4 & D- --> pi- pi- K+ & 1
  5 & D*+ --> pi+ D0 & 2
  6 & D0 --> K.L0 pi+ pi- & 5
}
```

Table 12 shows the resulting topology map. The results are the same as those displayed in the first two rows of Table 2.

Table 12: Signal decay trees and their respective initial-final states.

rowNo	signal decay tree (signal decay initial-final states)	iSigDcyTr	nEtr	nCEtr
1	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow e^+ \nu_e D^{*-} \gamma^F, \bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}, D^{*-} \rightarrow \pi^- \bar{D}^0,$ $D^{*+} \rightarrow \pi^+ D^0, \bar{D}^0 \rightarrow \pi^0 \pi^- K^+, D^0 \rightarrow \pi^0 \pi^+ K^-$ $(\Upsilon(4S) \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu \pi^0 \pi^+ \pi^+ \pi^- \pi^- K^+ K^- \gamma^F)$	0	3	3
2	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow \pi^0 \pi^+ \pi^+ \rho^- D^-, \bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}, \rho^- \rightarrow \pi^0 \pi^-,$ $D^- \rightarrow \pi^- \pi^- K^+, D^{*+} \rightarrow \pi^+ D^0, D^0 \rightarrow K_L^0 \pi^+ \pi^-$ $(\Upsilon(4S) \rightarrow \mu^- \bar{\nu}_\mu \pi^0 \pi^0 K_L^0 \pi^- \pi^+ \pi^+ \pi^- \pi^- \pi^- K^+)$	1	2	5

4.2. Decay initial-final states

In a few cases, we have interest in some decay initial-final states. Below is an example demonstrating the related item by taking the first two decay initial-final states listed in Table 3 as signals. Similar to IRA decay branches, decay initial-final states look like inclusive decay branches. Hence, except that only two columns are involved in the item, the format of the input to the item for decay initial-final states is identical to that for the component analysis over inclusive decay branches, which is introduced in Section 3.8. As we can see from the example, the numbers of identical particles are supported to be written in front of their textual names in order to simplify the textual expressions of the final states.

```
% Signal identification — decay initial-final states
{
  Y(4S) --> mu+ nu_mu 3 pi0 3 pi+ 4 pi- K+ K- & 2ndDcyIFStsInTb3
  Y(4S) --> 5 pi0 5 pi+ 5 pi- K+ K- & 2ndDcyIFStsInTb3
}
```

The obtained topology map is displayed in Table 13. The results are identical to those shown in the first two rows of Table 3.

Table 13: Signal decay initial-final states.

rowNo	signal decay initial-final states	iSigDcyIFSts2	nEtr	nCEtr
1	$\Upsilon(4S) \rightarrow \mu^+ \nu_\mu \pi^0 \pi^0 \pi^+ \pi^+ \pi^- \pi^- \pi^- \pi^- K^+ K^-$	0	18	18
2	$\Upsilon(4S) \rightarrow \pi^0 \pi^0 \pi^0 \pi^0 \pi^+ \pi^+ \pi^+ \pi^- \pi^- \pi^- \pi^- K^+ K^-$	1	18	36

4.3. Particles

Occasionally, we may want to identify some particles. The following example shows the associated item with the two particles D^{*+} and J/ψ set as signals. Except that only two columns are involved in the item, the format of the input to the item is identical to that for the component analysis over decay branches of particles, which is introduced in Section 3.3.

```
% Signal identification — particles
{
  D*+ Dsp
  J/psi Jpsi
}
```

Table 14 shows the resulting topology map. As a cross-check, the number of D^{*+} s in the table equals those in Tables 4, 5, and 6.

Table 14: Signal particles.

rowNo	signal particle	iSigP	nCase	nCCase
1	D^{*+}	0	45886	45886
2	J/ψ	1	2654	48540

4.4. Decay branches

On some occasions, we have to identify certain regular decay branches. Below is an example demonstrating the related item by taking the two decay branches $\bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}$ and $B^0 \rightarrow K_S^0 J/\psi$ as signals. Since regular decay branches also look like inclusive decay branches, except that only two columns are involved in the item, the format of the input to the item for regular decay branches is identical to that for the component analysis over inclusive decay branches,

629 which is introduced in Section 3.8.

```
630
631 % Signal identification — decay branches
632 {
633     B0 --> mu- anti-nu_mu D*+ & B2munuDsp
634     B0 --> K_S0 J/psi & B2KsJpsi
635 }
636
```

637 The obtained topology map is displayed Table 15. For the purpose of cross-checks, we note that
638 the number of $\bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}$ ($B^0 \rightarrow K_S^0 J/\psi$) in the table is equal to that in the first row of Table
639 5 (9).

Table 15: Signal decay branches.

rowNo	signal decay branch	iSigDcyBr	nCase	nCCase
1	$\bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}$	0	4154	4154
2	$B^0 \rightarrow K_S^0 J/\psi$	1	45	4199

640 4.5. Cascade decay branches

641 Sometimes, we are interested in certain cascade decay branches. The following example
642 shows the associated item with the two cascade decay branches $B^0 \rightarrow D^{*-} D_s^{*+}$, $D^{*-} \rightarrow \pi^- \bar{D}^0$,
643 $D_s^{*+} \rightarrow D_s^+ \gamma$ and $B^0 \rightarrow D^{*-} D_s^{*+}$, $D^{*-} \rightarrow \pi^- \bar{D}^0$ set as signals. While the first cascade decay
644 branch is identical to the fifth one in Table 7, the second is only part of it, which demonstrates
645 that the cascade decay branches supported in the item are not necessarily fully specified at the
646 level of certain hierarchy. Similar to decay trees, cascade decay branches are made up of regular
647 decay branches. Hence, the format of the input to the item for cascade decay branches is identical
648 to that for decay trees, which is introduced in Section 4.1.

```
649
650 % Signal identification — cascade decay branches
651 {
652     0 & B0 --> D*- D_s*+ & -1
653     1 & D*- --> pi- anti-D0 & 0
654     2 & D_s*+ --> D_s+ gamma & 0
655
656     0 & B0 --> D*- D_s*+ & -1
657     1 & D*- --> pi- anti-D0 & 0
658 }
659
```

660 Table 16 shows the resulting topology map. As a cross-check, the number of cases of the first
661 cascade decay branch in the table equals that of the fifth cascade decay branch in Table 7.

Table 16: Signal cascade decay branches.

rowNo	signal cascade decay branch	iSigCascDcyBr	nCase	nCCase
1	$B^0 \rightarrow D^{*-} D_s^{*+}, D^{*-} \rightarrow \pi^- \bar{D}^0, D_s^{*+} \rightarrow D_s^+ \gamma$	0	1119	1119
2	$B^0 \rightarrow D^{*-} D_s^{*+}, D^{*-} \rightarrow \pi^- \bar{D}^0$	1	1180	2299

662 4.6. Inclusive decay branches

663 In a few cases, we have to identify some inclusive decay branches. Below is an example
664 demonstrating the related item by taking the two inclusive decay branches $\bar{B}^0 \rightarrow D^{*+} + \text{anything}$
665 and $B^0 \rightarrow K_S^0 + \text{anything}$ as signals. Except that only two columns are involved in the item, the

format of the input to the item is identical to that for the component analysis over inclusive decay branches, which is introduced in Section 3.8.

```

668 % Signal identification — inclusive decay branches
669 {
670   anti-B0 --> D*+ & B2Dsp
671   B0 --> K_S0 & B2Ks
672 }
673
674

```

The obtained topology map is displayed in Table 17. As a cross-check, the number of $B^0 \rightarrow K_S^0 + \text{anything}$ in the table equals that in Table 9.

Table 17: Signal inclusive decay branches.

rowNo	signal inclusive decay branch	iSigIncDcyBr	nCase	nCCase
1	$\bar{B}^0 \rightarrow D^{*+} + \text{anything}$	0	41751	41751
2	$B^0 \rightarrow K_S^0 + \text{anything}$	1	153	41904

4.7. Inclusive cascade decay branches

Occasionally, we may have interest in certain inclusive cascade decay branches. The following example shows the associated item with the two inclusive cascade decay branches $\bar{B}^0 \rightarrow D^{*+} + \text{anything}$, $D^{*+} \rightarrow \pi^+ D^0$ and $B^0 \rightarrow K_S^0 J/\psi$, $K_S^0 \rightarrow \pi^+ \pi^-$, $J/\psi \rightarrow \mu^+ + \text{anything}$ set as signals. Similar to decay trees and cascade decay branches, inclusive cascade decay branches are made up of regular decay branches. Hence, the format of the input to the item for inclusive cascade decay branches is also identical to that for decay trees, which is introduced in Section 4.1. and the independent textual name “*” denotes anything.

```

685 % Signal identification — inclusive cascade decay branches
686 {
687   0 & anti-B0 --> D*+ * & -1
688   1 & D*+ --> pi+ D0 & 0
689
690   0 & B0 --> K_S0 J/psi & -1
691   1 & K_S0 --> pi+ pi- & 0
692   2 & J/psi --> mu+ * & 0
693 }
694
695

```

Table 18 shows the resulting topology map.

Table 18: Signal inclusive cascade decay branches.

rowNo	signal inclusive cascade decay branch	iSigIncCascDcyBrs	nCase	nCCase
1	$\bar{B}^0 \rightarrow D^{*+} + \text{anything}, D^{*+} \rightarrow \pi^+ D^0$	0	28367	28367
2	$B^0 \rightarrow K_S^0 J/\psi, K_S^0 \rightarrow \pi^+ \pi^-, J/\psi \rightarrow \mu^+ + \text{anything}$	1	1	28368

4.8. Intermediate-resonance-allowed decay branches

On some occasions, we need to identify certain IRA decay branches. Below is an example demonstrating the related item by taking the two IRA decay branches $D^{*+} \rightarrow \pi^0 \pi^+ \pi^- K^-$ and $J/\psi \rightarrow \pi^0 \pi^+ \pi^-$ as signals. Except that only two columns are involved in the item, the format of the input to the item is identical to that for the component analysis over IRA decay branches, which is introduced in Section 3.9.

```

704 % Signal identification — intermediate-resonance-allowed decay branches
705 {
706     D*+ --> K- pi+ pi+ pi0 & Dsp2K3Pi
707     J/psi --> pi+ pi- pi0 & Jpsi23Pi
708 }

```

709
710 The obtained topology map is displayed in Table 19. For the purpose of cross-checks, we
711 note that the number of $D^{*+} \rightarrow \pi^0 \pi^+ \pi^- K^-$ in the table is equal to that in Table 10.

Table 19: Signal IRA decay branches.

rowNo	signal IRA decay branch	iSigIRADcyBr	nCase	nCCase
1	$D^{*+} \rightarrow \pi^0 \pi^+ \pi^- K^-$	0	4971	4971
2	$J/\psi \rightarrow \pi^0 \pi^+ \pi^-$	1	59	5030

712 4.9. Essential topology tags

713 Table 20 summarizes and explains all of the essential topology tags involved in the signal
714 identification functionalities. For signal decay trees and signal decay initial-final states, there
715 are two sorts of topology tags. The first sort of tags, iSigDcyTr and iSigDcyIFSts, record the
716 default indices of the specified signal decay trees and signal decay initial-final states. They have
717 the similar interpretation as iDcyTr and iDcyIFSts, and are shown in the third columns of Tables
718 12 and 13. The second sort of tags, nameSigDcyTr and nameSigDcyIFSts, save the specified
719 aliases of the signal decay trees and signal decay initial-final states. In cases the aliases are not
720 specified, empty strings will be stored.

721 For the latter six kinds of signal identification, there are only one sort of topology tags, which
722 record the number of the specified particles or decay branches found in each entry. Similar to
723 the cases in the latter seven kinds of component analysis, in the topology tags, “i” in “_i” is the
724 default index of the specified particle or decay branch, and it ranges from 0 (included) to the
725 number of specified particles or decay branches (excluded). If the alias of the particle or decay
726 branch is also specified, the index “i” will be replaced with the alias.

Table 20: Essential topology tags involved in each kind of signal identification.

Signal identification kind	Topology tag	Interpretation
Decay trees	iSigDcyTr	index of signal decay tree
	nameSigDcyTr	name of signal decay tree
Decay initial-final states	iSigDcyIFSts	index of signal decay initial-final states
	nameSigDcyIFSts	name of signal decay initial-final states
Particles	nSigP_i	number of signal particle;s
Decay branches	nSigDcyBr_i	number of signal decay branch;es
Cascade decay branches	nSigCascDcyBr_i	number of signal cascade decay branch;es
Inclusive decay branches	nSigIncDcyBr_i	number of signal inclusive decay branch;es
Inclusive cascade decay branches	nSigIncCascDcyBr_i	number of signal inclusive cascade decay branch;es
IRA decay branches	nSigIRADcyBr_i	number of signal IRA decay branch;es

727 5. Common settings

728 From Sections 3 and 4, the optional parameters of the functionality items give us more choic-
729 es and thus help us do our jobs quicker and better. In addition to these parameters, many optional

items are designed and implemented to control the execution of the program in order to meet practical needs. Unlike the optional parameters, which only affect the individual functionalities to which they belong, the optional items have impact on all of the functionalities, or at least most of the functionalities. The current version of the program contains 24 commonly used items, which can be divided into the following three groups: items on the input of the program, items on the functionalities of the program, and items on the output of the program. This section introduces these items in the three groups, with each group in one subsection.

5.1. Settings on the input of the program

5.1.1. Input entries

The program normally processes all of the entries in the input samples, but sometimes only a part of the entries are needed to be (first) processed. Running the program over a big sample usually takes a long time. In such a case, it is a good habit to run the program first over a small part of the sample to check possible exceptions, and then over the whole sample if no exceptions are found or after the found exceptions are handled. Besides, a small number of entries is usually sufficient to do tests in the developing of the program. For these reasons, an item is developed to set up the maximum number of entries to be processed. Below is an example showing the item with the maximum number set at two thousand.

```
% Maximum number of entries to be processed
{
    2000
}
```

On some occasions, especially in the course of optimizing selection criteria, we need to run the program only over entries satisfying certain requirements. For this purpose, an item is developed to select entries. The following example shows the item with X set in the range $(-1, 1)$.

```
% Cut to select entries
{
    (X > -1) && (X < 1)
}
```

Notably, only a single-line selection requirement is supported in the item, like the cases in the methods Draw() and GetEntries() of the class TTree. In spite of this, such a requirement is able to express any requirement with the help of the parentheses “()” as well as the logical symbols “&&”, “||”, and “!”.

Occasionally, array variables are involved in the requirement. Under the circumstances, users have to tell the program how to determine the total logical value with the individual logical values. At present, two criteria are provided: (1) the total result is true as long as the result for one instance is true; (2) the total result is false as long as the result for one instance is false. By default, the second criterion is used in the program. One can alter it to the first one with the following item.

```
% Method to apply cut to array variables (Two options: T and F. Default: F)
{
    T
}
```

In the item, “T” and “F” stand for the first and second criteria, respectively.

5.1.2. Input decay branches

Normally, the program deals with all of the decay branches in every decay tree. However, examining all the branches is not always required in practice. Sometimes, we only concern the first n hierarchies of the branches. Similar to that in cascade decay branches of particles (as we introduce in Section 3.6), the hierarchy here reflects the rank of a decay branch in a decay tree. For example, in the decay tree $\Upsilon(4S) \rightarrow B^0 \bar{B}^0$, $B^0 \rightarrow e^+ \nu_e D^{*-} \gamma^F$, $\bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}$, $D^{*-} \rightarrow \pi^- \bar{D}^0$, $D^{*+} \rightarrow \pi^+ D^0$, $\bar{D}^0 \rightarrow \pi^0 \pi^- K^+$, $D^0 \rightarrow \pi^0 \pi^+ K^-$, the hierarchies of the seven individual branches are 1, 2, 2, 3, 3, 4, and 4, respectively. The program provides an item to set the maximum hierarchy. Below is an example showing the item with the maximum hierarchy set at one.

```
% Maximum hierarchy of heading decay branches to be processed in each event
{
  1
}
```

With the setting, the decay branches with hierarchy larger than one will be ignored by the program. For the component analysis over the decay trees of the $\Upsilon(4S)$ sample, only the first hierarchy of $\Upsilon(4S)$ decay branches are analyzed, and the result is shown in Table 21. From the table, not only $\Upsilon(4S) \rightarrow B^0 \bar{B}^0$ but also $\Upsilon(4S) \rightarrow B^0 B^0$ and $\Upsilon(4S) \rightarrow \bar{B}^0 \bar{B}^0$ are seen because of B^0 - \bar{B}^0 mixing.

Table 21: Decay trees and their respective initial-final states.

rowNo	decay tree (decay initial-final states)	iDcyTr	nEtr	nCEtr
1	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0$ ($\Upsilon(4S) \rightarrow B^0 \bar{B}^0$)	0	81057	81057
2	$\Upsilon(4S) \rightarrow B^0 B^0$ ($\Upsilon(4S) \rightarrow B^0 B^0$)	1	9487	90544
3	$\Upsilon(4S) \rightarrow \bar{B}^0 \bar{B}^0$ ($\Upsilon(4S) \rightarrow \bar{B}^0 \bar{B}^0$)	2	9456	100000

Similarly, in the case of the maximum hierarchy set at two, we could get the result of the component analysis over the first two hierarchies of $\Upsilon(4S)$ decay branches, as displayed in Table 22.

Table 22: Decay trees and their respective initial-final states.

rowNo	decay tree (decay initial-final states)	iDcyTr	nEtr	nCEtr
1	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow \mu^+ \nu_\mu D^{*-}, \bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}$ ($\Upsilon(4S) \rightarrow \mu^+ \mu^- \nu_\mu \bar{\nu}_\mu D^{*+} D^{*-}$)	936	136	136
2	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow e^+ \nu_e D^{*-}, \bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}$ ($\Upsilon(4S) \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu D^{*+} D^{*-}$)	1188	112	248
3	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow \mu^+ \nu_\mu D^{*-}, \bar{B}^0 \rightarrow e^- \bar{\nu}_e D^{*+}$ ($\Upsilon(4S) \rightarrow e^- \bar{\nu}_e \mu^+ \nu_\mu D^{*+} D^{*-}$)	268	110	358
4	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow D^{*-} D_s^{*+}, \bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}$ ($\Upsilon(4S) \rightarrow \mu^- \bar{\nu}_\mu D^{*+} D^{*-} D_s^{*+}$)	2063	72	430
5	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow e^+ \nu_e D^{*-}, \bar{B}^0 \rightarrow e^- \bar{\nu}_e D^{*+}$ ($\Upsilon(4S) \rightarrow e^+ e^- \nu_e \bar{\nu}_e D^{*+} D^{*-}$)	95	71	501
rest	$\Upsilon(4S) \rightarrow \text{others (81609 in total)}$ ($\Upsilon(4S) \rightarrow \text{corresponding to others}$)	—	99499	100000

Sometimes, we do not care about the decay of some particles. One can make the program ignore their decay branches with the following item. With the setting in the example, the decay of B^0 and \bar{B}^0 will be ignored by the program.

```

805
806 % Ignore the decay of the following particles
807 {
808     B0
809     anti-B0
810 }
811

```

At some other times, we have interest in the decay of some particles but not in the decay of their daughters. To handle this case, the following item is developed to make the program ignore the decay of their daughters. In the following example, the decay of the daughters of B^0 and \bar{B}^0 will be ignored by the program.

```

816
817 % Ignore the decay of the daughters of the following particles
818 {
819     B0
820     anti-B0
821 }
822

```

The two settings above have the same effects as those in the previous paragraph which set the maximum hierarchy at one and two, and hence the corresponding results are identical to those shown in Tables 21 and 22.

As mentioned in Section 2.4, the decay $\pi^0 \rightarrow \gamma\gamma$ is ignored by default. On the occasions when we need to identify the signals involving the decay, we can make the program retain the decay with the item below set to “Y”.

```

829
830 % Retain the decay of pi0 to gamma gamma (Two options: Y and N. Default: N)
831 {
832     Y
833 }
834

```

Besides, if needed, one can make the program ignore other final decay branches, such as $\eta \rightarrow \gamma\gamma$ and $K_S^0 \rightarrow \pi^+\pi^-$, with the following item.

```

837
838 % Ignore the following final decay branches
839 {
840     eta --> gamma gamma
841     K_S0 --> pi+ pi-
842 }
843

```

5.1.3. Initial and final state radiation photons

Initial state radiation (ISR) and final state radiation (FSR) are inevitable physical effects in e^+e^- colliding experiments. Therefore, ISR and FSR photons are often involved in the inclusive MC samples. The program processes them together with other particles in the default case. To distinguish them from other photons, the program tries to label them in the output txt, tex and pdf files. Sometimes, these photons are marked out beforehand with special PDG codes according to particle status information from generators. One can inform the program of these PDG codes by the following two items.

```

851
852 % PDG code of ISR photons (Default: 222222222)
853 {
854     222222222
855 }
856

```



```

856
857
858 % PDG code of FSR photons (Default: -22)
859 {
860     -22
861 }

```

862
863 In this case, the program is able to label the ISR and FSR photons as γ^i (gammai) and γ^f (gammaf) in the output pdf (txt) files, respectively.
864

865 On other occasions, ISR and FSR photons are not marked out in advance due to some reasons.
866 In such cases, the program have to identify them by itself according to the following rules:
867 photons which have no mothers recorded in the arrays of the PDG codes and mother indices
868 are considered as generalized ISR photons, while other photons which have at least one e^\pm , μ^\pm ,
869 π^\pm , K^\pm , p , or \bar{p} sister are taken as generalized FSR photons. Here, the modifier “generalized”
870 is used because the rules can not determine the types of the photons in absolute accuracy. For
871 example, photons from radiative decays might be mistaken as ISR and FSR photons. Despite this,
872 generalized ISR and FSR photons are good concepts, particularly in cases where the sources of
873 the photons are not required to be distinguished clearly. The program will label the generalized
874 ISR and FSR photons as γ^I (gammaI) and γ^F (gammaF) in the output pdf (txt) files, respectively.

875 Notably, we are not concerned about these ISR and FSR photons in many cases. If they have
876 already been marked out beforehand, one can make the program ignore them accurately by set-
877 ting the following two items to “Ys”.

```

878
879 % Ignore ISR photons (Three options: Ys, Yg and N. Default: N)
880 {
881     Ys
882 }
883
884
885 % Ignore FSR photons (Three options: Ys, Yg and N. Default: N)
886 {
887     Ys
888 }

```

889
890 In cases that these photons are not marked in advance, the option “Yg” can be used to ignore the
891 generalized ISR and FSR photons. In “Ys” and “Yg”, “s” and “g” are the initials of the words
892 “strict” and “generalized”, respectively.

893 5.2. Settings on the functionalities of the program

894 5.2.1. Candidate based analysis

895 According to the number of signal candidates in an event that are selected and retained to
896 extract physics results, data analysis in high energy experiments can be divided into the following
897 two categories: event based analysis and candidate based analysis. While at most one candidate
898 in an event is kept in event based analysis, one or more candidates in an event can be retained in
899 candidate based analysis. Generally, the quantities related to a candidate are stored in an entry of
900 the TTree objects in the root files. Thus, one or more entries relate to an event in candidate based
901 analysis, while only one entry corresponds to an event in event based analysis. Normally, the
902 indices of candidates within an event are stored in the corresponding entries in candidate based
903 analysis.

904 By default, the program analyzes the input entries one by one. In this case, the events with
905 multiple candidates will be processed repeatedly. Particularly, the number of physics processes

at the sample level will be over counted. One can make the program avoid the problem by inputting “Y” to the following item.

```

909      % Avoid over counting for candidate based analysis (Two options: Y and N. Default: N)
910      {
911          Y
912      }

```

Also, the indices of candidates within an event are required. We can tell the program the related TBranch name with the following item.

```

917      % TBranch name of the indices of candidates in an event (Default: __candidate__)
918      {
919          iCandidate
920      }

```

With the settings, the program will process the first entry of each event in a normal way, including obtaining and storing the topology tags; it will not analyze the other entries of the same event, but only store the same topology tags to them.

5.2.2. Charge conjugation

Charge conjugation is an important concept in high energy physics. By default, charge conjugate objects (particles and decays) are processed separately in the program. However, we need to handle them together in many physics studies because of the sameness between them. One can have the program process them together with the item below set to “Y”.

```

931      % Process charge conjugate objects together (Two options: Y and N. Default: N)
932      {
933          Y
934      }

```

Performing topology analysis with the setting inserts new topology tags in the output root files and adds new counters to topology maps in the output txt, tex, and pdf files. Tables 23 and 24 list and interpret all of the topology tags related to charge conjugation involved in the component analysis and signal identification functionalities, respectively.

As an example, we perform the component analysis over decay trees with the charge conjugate item. Table 25 shows the obtained topology map. Besides the columns in Table 2, two additional columns with the headers “nCCEtr” and “nAllEtr” are inserted in the table. Here, “nCCEtr” represents the number of entries involving the charge conjugate decay trees, and “nAllEtr” is the sum of “nEtr” and “nCCEtr”. In addition to “iDcyTr”, “iCCDcyTr” is also inserted in the output root files as a topology tag. It is short for charge conjugate index of decay tree. For self-charge-conjugate decay trees, it has the value 0; for non-self-charge-conjugate decay trees, it has the value 1 or -1: while 1 tags the decay trees listed in the topology maps, -1 indicates their charge conjugate decay trees. Whereas the values of “iDcyTr” for each decay tree and its charge conjugate decay tree are equal in order to indicate their sameness, the values of “iCCDcyTr” for them are opposite so as to reflect their difference.

As another example, we carry out the component analysis over the decay branches of D^{*+} . The resulting topology map of D^{*+} is displayed in Table 26. Compared with Table 4, two new columns are added to the table, and their headers “nCCEtr” and “nAllEtr” have the similar meanings as “nCCEtr” and “nAllEtr” in Table 25. For a specified particle, what we want to further record with topology tags are as follows: (1) whether it is self-charge-conjugate; (2) whether its

Table 23: Topology tags related to charge conjugation involved in each kind of component analysis. For the latter seven kinds of component analysis, the topology tags in the 1) and 2) groups are only designed for the self-charge-conjugate and non-self-charge-conjugate particles and decay branches, respectively. The acronyms “cc” and index_{cc} are short for “charge conjugate” and “charge conjugate index”, respectively.

Component analysis kind	Topology tag	Interpretation
Decay trees	iCcDcyTr	index_{cc} of decay tree
Decay initial-final states	iCcDcyIFSts	index_{cc} of decay initial-final states
Decay branches of particles	iCcPDcyBr.i	index_{cc} of particle _i
	1) iCcDcyBrP.i.j	index_{cc} of decay branch of the j th particle _i
	2) nCcPDcyBr.i	number of cc particle _i s (decay branches)
	2) iDcyBrCcP.i.j	index of decay branch of the j th cc particle _i
	2) nAllPDcyBr.i	number of all particle _i s (decay branches)
Production branches of particles	iCcPProdBr.i	index_{cc} of particle _i
	1) iCcProdBrP.i.j	index_{cc} of production branch of the j th particle _i
	2) nCcPProdBr.i	number of cc particle _i s (production branches)
	2) iProdBrCcP.i.j	index of production branch of the j th cc particle _i
	2) nAllPProdBr.i	number of all particle _i s (production branches)
Mothers of particles	iCcPMoth.i	index_{cc} of particle _i
	1) iCcMothP.i.j	index_{cc} of mother of the j th particle _i
	2) nCcPMoth.i	number of cc particle _i s (mothers)
	2) PDGMothCcP.i.j	PDG code of mother of the j th cc particle _i
	2) nAllPMoth.i	number of all particle _i s (mothers)
Cascade decay branches of particles	iCcPCascDcyBr.i	index_{cc} of particle _i
	1) iCcCascDcyBrP.i.j	index_{cc} of cascade decay branch of the j th particle _i
	2) nCcPCascDcyBr.i	number of cc particle _i s (cascade decay branches)
	2) iCascDcyBrCcP.i.j	index of cascade decay branch of the j th cc particle _i
	2) nAllPCascDcyBr.i	number of all particle _i s (cascade decay branches)
Decay final states of particles	iCcPDcyFSt.i	index_{cc} of particle _i
	1) iCcDcyFStP.i.j	index_{cc} of decay final state of the j th particle _i
	2) nCcPDcyFSt.i	number of cc particle _i s (decay final states)
	2) iDcyFStCcP.i.j	index of decay final state of the j th cc particle _i
	2) nAllPDcyFSt.i	number of all particle _i s (decay final states)
Inclusive decay branches	iCcIncDcyBr.i	index_{cc} of inclusive decay branch _i
	1) iCcDcyBrIncDcyBr.i.j	index_{cc} of decay branch of the j th inclusive decay branch _i
	2) nCcIncDcyBr.i	number of cc inclusive decay branch _i es
	2) iDcyBrCcIncDcyBr.i.j	index of decay branch of the j th cc inclusive decay branch _i
	2) nAllIncDcyBr.i	number of all inclusive decay branch _i es
IRA decay branches	iCcIRADcyBr.i	index_{cc} of IRA decay branch _i
	1) iCcDcyBrIRADcyBr.i.j	index_{cc} of decay branch of the j th IRA decay branch _i
	2) nCcIRADcyBr.i	number of cc IRA decay branch _i es
	2) iDcyBrCcIRADcyBr.i.j	index of decay branch of the j th cc IRA decay branch _i
	2) nAllIRADcyBr.i	number of all IRA decay branch _i es

Table 24: Topology tags related to charge conjugation involved in each kind of signal identification. For the latter six kinds of signal identification, the topology tags in the *) groups are only designed for the non-self-charge-conjugate particles and decay branches. The acronyms “cc” and index_{cc} are short for “charge conjugate” and “charge conjugate index”, respectively.

Signal identification kind	Topology tag	Interpretation
Decay trees	iCcSigDcyTr	index_{cc} of signal decay tree
Decay initial-final states	iCcSigDcyIFSts	index_{cc} of signal decay initial-final states
Particles	iCcSigP.i	index_{cc} of signal particle _i
	*) nCcSigP.i	number of cc signal particle _i s
	*) nAllSigP.i	number of all signal particle _i s
Decay branches	iCcSigDcyBr.i	index_{cc} of signal decay branch _i
	*) nCcSigDcyBr.i	number of cc signal decay branch _i es
	*) nAllSigDcyBr.i	number of all signal decay branch _i es
Cascade decay branches	iCcSigCascDcyBr.i	index_{cc} of signal cascade decay branch _i
	*) nCcSigCascDcyBr.i	number of cc signal cascade decay branch _i es
	*) nAllSigCascDcyBr.i	number of all signal cascade decay branch _i es
Inclusive decay branches	iCcSigIncDcyBr.i	index_{cc} of signal inclusive decay branch _i
	*) nCcSigIncDcyBr.i	number of cc signal inclusive decay branch _i es
	*) nAllSigIncDcyBr.i	number of all signal inclusive decay branch _i es
Inclusive cascade decay branches	iCcSigIncCascDcyBr.i	index_{cc} of signal inclusive cascade decay branch _i
	*) nCcSigIncCascDcyBr.i	number of cc signal inclusive cascade decay branch _i es
	*) nAllSigIncCascDcyBr.i	number of all signal inclusive cascade decay branch _i es
IRA decay branches	iCcSigIRADcyBr.i	index_{cc} of signal IRA decay branch _i
	*) nCcSigIRADcyBr.i	number of cc signal IRA decay branch _i es
	*) nAllSigIRADcyBr.i	number of all signal IRA decay branch _i es

956 decay branches are self-charge-conjugate, if it is self-charge-conjugate; (3) the number and the
957 indices of the decay branches of its charge-conjugate particle, if it is not self-charge-conjugate.

Table 25: Decay trees and their respective initial-final states.

rowNo	decay tree (decay initial-final states)	iDcyTr	nEtr	nCcEtr	nAllEtr	nCEtr
1	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow e^+ \nu_e D^{*-} \gamma^F, \bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}, D^{*-} \rightarrow \pi^- \bar{D}^0,$ $D^{*+} \rightarrow \pi^+ D^0, \bar{D}^0 \rightarrow \pi^0 \pi^- K^+, D^0 \rightarrow \pi^0 \pi^+ K^-$ $(\Upsilon(4S) \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu \pi^0 \pi^+ \pi^- \pi^- K^+ K^- \gamma^F)$	20870	3	0	3	3
2	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow \mu^+ \nu_\mu D^-, \bar{B}^0 \rightarrow e^- \bar{\nu}_e D^{*+}, D^- \rightarrow e^- \bar{\nu}_e \pi^- K^+,$ $D^{*+} \rightarrow \pi^+ D^0, D^0 \rightarrow \pi^0 \pi^+ K^-$ $(\Upsilon(4S) \rightarrow e^- e^+ \bar{\nu}_e \nu_e \mu^+ \nu_\mu \pi^0 \pi^+ \pi^- \pi^- K^+ K^-)$	3722	1	1	2	5
3	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow \pi^0 \pi^+ \pi^- \rho^- D^-, \bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}, \rho^- \rightarrow \pi^0 \pi^-,$ $D^- \rightarrow \pi^- \pi^- K^+, D^{*+} \rightarrow \pi^+ D^0, D^0 \rightarrow K_S^0 \pi^+ \pi^-$ $(\Upsilon(4S) \rightarrow \mu^- \bar{\nu}_\mu \pi^0 \pi^0 K_S^0 \pi^+ \pi^+ \pi^- \pi^- \pi^- K^+ K^-)$	5295	2	0	2	7
4	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow e^+ \nu_e D^{*-} \gamma^F, \bar{B}^0 \rightarrow \pi^0 \pi^+ \pi^- \pi^- D^{*+},$ $D^{*-} \rightarrow \pi^0 D^-, D^{*+} \rightarrow \pi^+ D^0, D^- \rightarrow \pi^- \pi^- K^+, D^0 \rightarrow \pi^0 \pi^+ K^-$ $(\Upsilon(4S) \rightarrow e^+ \nu_e \pi^0 \pi^0 \pi^+ \pi^+ \pi^- \pi^- \pi^- K^+ K^- \gamma^F)$	10206	1	1	2	9
5	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow \mu^+ \nu_\mu D^{*-}, \bar{B}^0 \rightarrow \mu^+ \nu_\mu D^{*+}, D^{*-} \rightarrow \pi^0 D^-,$ $D^{*+} \rightarrow \pi^- \bar{D}^0, D^- \rightarrow \pi^0 \pi^- K_S^0, \bar{D}^0 \rightarrow \pi^0 \pi^- K^+, K_S^0 \rightarrow \pi^+ \pi^-$ $(\Upsilon(4S) \rightarrow \mu^+ \mu^+ \nu_\mu \nu_\mu \pi^0 \pi^0 \pi^+ \pi^- \pi^- \pi^- K^+ K^-)$	11916	1	1	2	11
rest	$\Upsilon(4S) \rightarrow \text{others (99969 in total)}$ $(\Upsilon(4S) \rightarrow \text{corresponding to others})$	—	—	—	99989	100000

Hence, in addition to “nPDcyBr.i” and “iDcyBrP.i.j”, the following topology tags are also inserted in the output root files: “iCCPDcyBr.i” for all specified particles; “iCCDcyBrP.i.j” for self-charge-conjugate particles only; and “nCCPDcyBr.i”, “iDcyBrCCP.i.j”, and “nAllPDcyBr.i” for non-self-charge-conjugate particles only. Here, “iCCPDcyBr.i” is short for charge conjugate index of the i^{th} particle specified for its decay branches. For self-charge-conjugate particles, it has the value 0; for non-self-charge-conjugate particles, it has the value 1.

The topology tag “iCCDcyBrP.i.j” denotes charge conjugate index of decay branch of the j^{th} instance of the i^{th} particle. It is to “iDcyBrP.i.j” what “iCCDcyTr” is to “iDcyTr”. The topology tag “iDcyBrCCP.i.j” has the similar meaning as “iDcyBrP.i.j”, but it is designed for the charge conjugate particle of the i^{th} particle. Particularly, it ranges from 0 (included) to the number of the categories of decay branches of the i^{th} particle found in the sample (excluded). The topology tag “nCCPDcyBr.i” stands for the number of the charge conjugate i^{th} particles (or their decay branches) found in each entry, and “nAllPDcyBr.i” is the sum of “nPDcyBr.i” and “nCCPDcyBr.i”.

Table 26: Decay branches of D^{*+} .

rowNo	decay branch of D^{*+}	iDcyBrP	nCase	nCCCase	nAllCase	nCCCase
1	$D^{*+} \rightarrow \pi^+ D^0$	0	31180	31291	62471	62471
2	$D^{*+} \rightarrow \pi^0 D^+$	1	13978	14166	28144	90615
3	$D^{*+} \rightarrow D^+ \gamma$	2	700	721	1421	92036
4	$D^{*+} \rightarrow \pi^+ D^0 \gamma^F$	3	28	36	64	92100
5	$D^{*+} \rightarrow \pi^0 D^+ \gamma$	4	0	1	1	92101

5.2.3. Settings only on signal identification

Normally, the signals specified in the signal identification functionality items are both tagged and counted by executing the program one time. In the case of a huge sample that will take a long time, it is a good idea to first tag the signals with multiple jobs each running on one machine, and then count the tagged signals together. One can make the program carry out the idea by setting the following item to “T” and “C” in the first and second steps, respectively. Here, “T” and “C” stand for tagging and counting, respectively.

```
% Analysis tasks for signal identifications (Three options: TC, T and C. Default: TC)
{
  T
}
```

By default, the signals set in the signal identification functionality items are listed in the output txt, tex, and pdf files in the sequence they are specified. In cases of plenty of signals, there is probably a need to sort them according to the number of cases found in the input samples. One can have the program do the sorting by inputting “Y” to the item below.

```
% Sort the signals in the topology maps related to signal identifications (Two options: Y and N. Default: N)
{
  Y
}
```

5.3. Settings on the output of the program

By default, decay objects (trees, initial-final states, and branches) are left-aligned in the output pdf files. If one likes it, he/she can request the program to center them by setting the following item to “Y”.

```

999      % Center decay objects in output pdf files (Two options: Y and N. Default: N)
1000      {
1001          Y
1002      }

```

As mentioned in Section 2.4, after the execution of the program, one or more root files will be output to save topology tags. By default, the program switches to a new output file whenever the size of the TTree object in memory exceeds 3 GB. In addition to this, the program provides an item to control the switch of output files by setting the maximum number of entries to be saved in a single output file. The following example shows the item with the maximum number set to 1 million.

```

1010
1011      % Maximum number of entries to be saved in a single output root file
1012      {
1013          1000000
1014      }

```

Besides, one can have the program generate one output file by one input file with the following item set to “Y”.

```

1018
1019      % One output root file by one input root file (Two options: Y and N. Default: N)
1020      {
1021          Y
1022      }

```

In default cases, flat TBranches are used to store topology tags in the output root files. This is necessary for the Belle II experiment, as array TBranches are not recommended to use in physics analyses in order to use other tools such as NumPy [10] and pandas [11]. However, since array TBranches are elegant and efficient in organizing and storing homogeneous data, sometimes it is better to use them than flat TBranches in other experiments, such as the BESIII experiment. One can make the program use array TBranches to store topology tags by inputting “Y” to the item below.

```

1031
1032      % Use array tbranches to store topology tags in output root files when possible (Two options: Y and N. Default: N)
1033      {
1034          Y
1035      }

```

By default, to facilitate the validation of topology analysis results, the input TBranches are copied to the output root files along with other TBranches for physics analyses. However, they often occupy too much disk space and are useless for following physics analyses. In the case of being flat, a massive amount of these TBranches also looks awkward. Thus, after the validation with a small sample, it would be better to remove these TBranches from the output root files. One can request the program to perform this removal operation before it terminates by setting the following item to “Y”.

```

1044
1045      % Remove the input tbranches from output root files (Two options: Y and N. Default: N)
1046      {
1047          Y
1048      }

```

1049 6. Auxiliary facilities

This section introduces some auxiliary facilities for the use of the program, including a card file to preset frequently used items and two commands implemented in tex source files. Different from that presented in the previous four sections, the content presented in this section is not the

essential part of the program. However, with these auxiliary facilities, we can make the program do our jobs better and quicker on some occasions.

6.1. The underlying card file

A card file, namely “underlying_topoana.card” under the share subdirectory, to preset frequently used items is developed to assist the card file specified by the first argument of the command “topoana.exe”. Here, we refer to the former and latter card files as underlying and primary, respectively. In general, the primary card file is sufficient to set items for the execution of the program. However, considering some items are frequently used with constant inputs by a user or a group of users, it is better to move the items from the primary card file to the underlying card file, in order to make the primary card file more concise and make us more focused on the items specially set for the dedicated topology analysis.

One can decide whether to set an item in the underlying card file according to his/her own needs. Here, we introduce some frequently used items that are suitable to be put in the underlying card file as follows. As mentioned in Section 2.3, the items related to the storage type and TBranches names of the input data are usually fixed for a user or a group of users. Thus, it is quite appropriate to move them to the underlying card file. We have to process charge conjugation particles and decays together in many physics studies. In such studies, it is also a good practice to put the item on charge conjugation in the underlying card file.

The program first reads the items in the underlying card file and then reads those in the primary card file. The items set in the underlying card file can be reset in the primary card file. In such a case, the inputs in the underlying card file will be replaced by their counterparts in the primary card file.

6.2. Commands implemented in tex source files

The output pdf files can be checked after the execution of the program. If their styles are not to our taste, we can edit the corresponding tex source files to get the desired styles, according to the regular LaTeX rules. Besides the rules, two commands are implemented in the tex source files to help us edit the files quickly and easily for two common desired styles.

By default, topology tags are listed along with topology maps in the output txt, tex, and pdf files. However, only the topology maps are needed on some occasions, especially in presentations. In such cases, one can suppress the topology tags in the output tex and pdf files by simply changing the definition of the `cmtTopoTags` command from the nominal one

```
\newcommand{\topoTags}[1]{#1}
```

to the alternative one

```
\newcommand{\topoTags}[1]{} 
```

in the preamble of the text source files. Here, “#1” is the formal parameter of the string for the topology tags. With the nominal definition, “\topoTags{#1}” returns the string exactly, while with the alternative definition it only returns an empty string. That is why the definition below is able to suppress the topology tags.

After the revision of the tex source files, one can re-compile them with the `pdflatex` command. Usually, the `pdflatex` command has to be executed two or three times for a fully compiled pdf file,

and many undesired files in other formats are generated during the compilation. To execute the `pdflatex` command and remove the undesired files at one stroke, we develop a bash script, namely “`getPdfFlFromTexFl.sh`” under the “utilities” subdirectory. The script should be executed with the following command line: `getPdfFlFromTexFl.sh texFileName`. Compiling the tex source files with the script is recommended.

7. Summary

We develop a program, namely `TopoAna`, with C++, ROOT and LaTeX for the topology analysis of inclusive MC samples in high energy physics experiments at e^+e^- colliders. This paper provides an essential description of the program, including a basic introduction of the program, two sorts of functionalities of the program — component analysis and signal identification, and some common settings and auxiliary facilities for the executing of the program.

Since it does not rely on any specific software frameworks, the program applies to many high energy physics experiments. Up to now, it has been put into use in three experiments at e^+e^- colliders: the BESIII, Belle, and Belle II experiments. Besides these experiments, it can also be used in the PANDA experiment [13], which is an anti-proton annihilation experiment under construction at Darmstadt, Germany. In addition, the program is also applicable to the pre-research of future e^+e^- colliding experiments, such as the circular electron positron collider (CEPC) [14, 15] experiment in China, the super charm-tau factory (SCTF) experiment [16] in Russia, and the super tau-charm factory (STCF) experiment [17] in China. These experiments offer a wide range of potential uses of the program. With more user needs coming out in the future, we will further extend and perfect it to make it more powerful and well-rounded.

Acknowledgements

This project is supported by the National Natural Science Foundation of China under Grants No. 11575017, No. 11661141008, No. 11761141009, and No. 11975076; the CAS Center for Excellence in Particle Physics (CCEPP). In addition, we would like to thank all of the people who have helped us in the development of the program. We first thank Prof. Changzheng Yuan, Bo Xin, and Haixuan Chen for their help at the early stage of developing the program. We are particularly grateful to Prof. Xingtao Huang for his comments on the principles and styles of the program, to Remco de Boer for his suggestions on the tex output and the use of GitHub, and to Xi Chen for his discussions on the core algorithms. We are especially indebted to Prof. Xiqing Hao, Longke Li, Xiaoping Qin, Ilya Komarov, Yubo Li, Guanda Gong, Suxian Li, Junhao Yin, Prof. Xiaolong Wang, and Yeqi Chen for their advice in extending and perfecting the program. Also, we thank Xi’an Xiong, Runqiu Ma, Wencheng Yan, Sen Jia, Hongpeng Wang, Jiawei Zhang, Hongrong Qi, Jiajun Liu, Maoqiang Jing, Yi Zhang, Wei Shan, and Yadi Wang for their efforts in helping us test the program.

References

- [1] ROOT User’s Guide, Available online: <https://root.cern/root/html/doc/guides/users-guide/ROOTUsersGuide.html>.
- [2] K.T. Chao, Y.F. Wang, et al., Physics at BES-III, *Int. J. Mod. Phys. A* 24 (2009) S1-794.
- [3] M. Ablikim, et al. (BESIII Collaboration), White Paper on the Future Physics Programme of BESIII, [arXiv:1912.05983](https://arxiv.org/abs/1912.05983).
- [4] E. Kou, et al., *Prog. Theor. Exp. Phys.* 2019 (2019) 123C01.

- 1138 [5] J. Brodzicka, T. Browder, P. Chang, et al., Prog. Theor. Exp. Phys. 2012 (2012) 04D001.
- 1139 [6] Documentation of the TFile class, Available online: <https://root.cern/root/html534/TFile.html>.
- 1140 [7] Documentation of the TTree class, Available online: <https://root.cern/root/html534/TTree.html>.
- 1141 [8] M. Tanabashi, et al. (Particle Data Group), Phys. Rev. D 98 (2018) 030001.
- 1142 [9] Documentation of the TBranch class, Available online: <https://root.cern/root/html534/TBranch.html>.
- 1143 [10] Documentation of NumPy, Available online: <https://numpy.org/devdocs/>.
- 1144 [11] Documentation of pandas, Available online: <https://pandas.pydata.org/pandas-docs/stable/>.
- 1145 [12] Documentation of the TChain class, Available online: <https://root.cern/root/html534/TChain.html>.
- 1146 [13] W. Erni, et al. (PANDA Collaboration), Physics Performance Report for PANDA: Strong Interaction Studies with
1147 Antiprotons, [arXiv:0903.3905](https://arxiv.org/abs/0903.3905).
- 1148 [14] CEPC CDR Volume 1 (Accelerator), Available online: http://cepc.ihep.ac.cn/CEPC_CDR_Vol1_Accelerator.pdf.
- 1149 [15] CEPC CDR Volume 2 (Physics & Detector), Available online: [http://cepc.ihep.ac.cn/CEPC_CDR_Vol2_Physics-](http://cepc.ihep.ac.cn/CEPC_CDR_Vol2_Physics-Detector.pdf)
1150 [Detector.pdf](http://cepc.ihep.ac.cn/CEPC_CDR_Vol2_Physics-Detector.pdf).
- 1151 [16] A.E. Bondar, et al. (Charm-Tau Factory Collaboration), Phys. Atom. Nucl. 76 (2013) 1072.
- 1152 [17] Q. Luo, D. Xu, “Progress on Preliminary Conceptual Study of HIEPA, a Super Tau-Charm Factory in China”, in
1153 Proc. 9th International Particle Accelerator Conf. (IPAC2018), Vancouver, BC, Canada, 422.