

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

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

^aSchool of Physics, Beihang University, Beijing 100191, China

^bSchool of Physics and Microelectronics, Zhengzhou University, Zhengzhou 450000, China

^cInstitute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China

^dKey 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 physics processes involved in the samples, including the types of processes and the number of processes in each type, is a great help to investigating signals and backgrounds. To help analysts obtain the physics process information from the truth information of the samples, we develop a physics process analysis program, TopoAna, with C++, ROOT, and LaTeX. The program implements the functionalities of component analysis and signal identification with many kinds of fine, customizable classification and matching algorithms. It tags physics processes in individual events accurately in the output root files, and exports the physics process information at the sample level clearly to the output plain text, tex source, and pdf files. 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 similar experiments is also prospective.

Keywords: event type; 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/generic Monte-Carlo (MC) samples 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

*Corresponding author.

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

The program is now available at <https://github.com/buaazhouxingyu/topoana>.

Preprint submitted to Computer Physics Communications

June 12, 2020

23 optimization of preliminary criteria is often needed in the process. Under the circumstances, a
 24 comprehensive understanding of the samples is required. In particular, a clear knowledge of the
 25 physics processes, or event types, involved in the samples is quite helpful. To be specific, the
 26 physics process information includes the types of processes and the number of processes in each
 27 type, involved both in the entire samples and in the individual events. Here, the physics process
 28 could be a complete production and decay process involved in an event, or merely a part of it,
 29 such as the decay of an intermediate resonance. With the information, one can figure out the
 30 main backgrounds (especially the peaking ones), and optimize the selection criteria further by
 31 analyzing the differences between the main backgrounds and the signals. Even if it is difficult
 32 to further suppress these backgrounds, the knowledge of their types is beneficial to estimate the
 33 systematic uncertainties associated with them.

34 The analysis of the physics process information described above is a sort of component anal-
 35 ysis. It is complex since it has to classify physics processes actively and finely. Another sort
 36 of physics process analysis often required in practice is signal identification, which only aims
 37 to search for certain processes of interests. It is relatively simple because its core technique is
 38 merely pattern matching. Mostly, signal and background events coexist in inclusive MC samples.
 39 It is useful to differentiate them in such cases. The identified signal events can be used to make
 40 up a signal sample in the absence of specialized signal samples, or they can be removed to avoid
 41 repetition in the presence of specialized signal samples. Occasionally, we have to pick out some
 42 decay branches in order to re-weight them according to new theoretical predictions or updated
 43 experimental measurements. Signal identification also plays a part in this occasion.

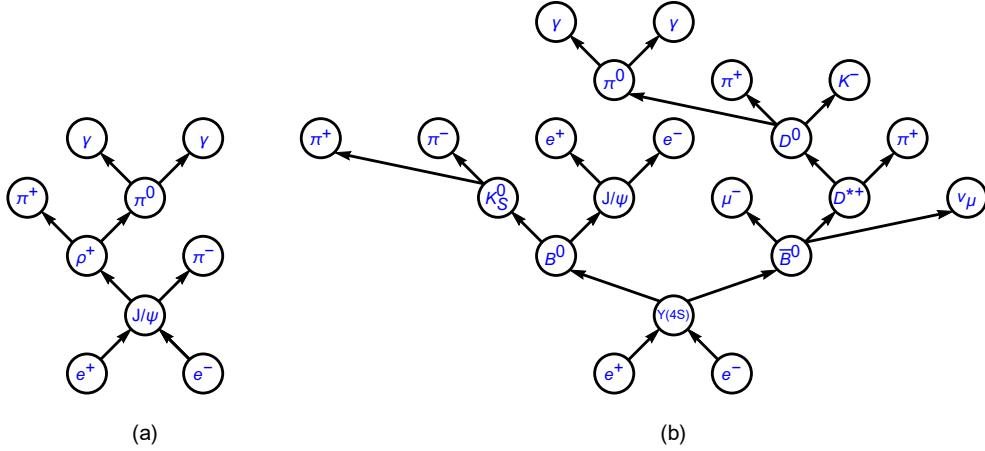


Figure 1: Topology diagrams of (a) $e^+e^- \rightarrow J/\psi$, $J/\psi \rightarrow \rho^+\pi^-$, $\rho^+ \rightarrow \pi^+\pi^0$, $\pi^0 \rightarrow \gamma\gamma$ and (b) $e^+e^- \rightarrow \Upsilon(4S)$, $\Upsilon(4S) \rightarrow B^0\bar{B}^0$, $B^0 \rightarrow K_S^0 J/\psi$, $\bar{B}^0 \rightarrow \mu^- D^{*+} \nu_\mu$, $K_S^0 \rightarrow \pi^+\pi^-$, $J/\psi \rightarrow e^+e^-$, $D^{*+} \rightarrow D^0\pi^+$, $D^0 \rightarrow \pi^0\pi^+K^-$, $\pi^0 \rightarrow \gamma\gamma$. As if trees grow, the diagrams are plotted from bottom to top.

44 Processes in high energy physics can be visualized with topology diagrams. As an example,
 45 Fig. 1 shows the topology diagrams of two typical physics processes occurring at e^+e^- colliders.
 46 From the figure, the hierarchies of the processes and the relationships among the particles are
 47 clearly illustrated with the diagrams. Though the complexities of topology diagrams vary with
 48 physics processes, there is only one diagram corresponding to each process. For this reason, we
 49 refer to the physics process information/analysis mentioned thereinbefore as topology informa-

50 tion/analysis hereinafter. The component analysis and signal identification introduced above are
51 exactly the two categories of topology analysis that will be discussed in this paper.

52 Since the raw topology truth information of inclusive MC samples is counter-intuitive, di-
53 verse, and overwhelming, it is difficult for analysts to check the topology information of the
54 samples directly. To help them do the checks quickly and easily, a topology analysis program
55 called TopoAna is developed with C++, ROOT [1], and LaTeX. Here, C++ is the programming
56 language, ROOT is the C++ based data analysis software universally used in modern high energy
57 physics experiments, and LaTeX is used for generating pdf documents containing the obtained
58 topology information. The program implements the functionalities of component analysis and
59 signal identification based on accurate pattern matching. To meet a variety of practical require-
60 ments, many kinds of fine, customizable classification and matching algorithms are implemented
61 in the program. Generally, the program recognizes, categorizes, and counts physics processes in
62 each event in the samples, and tags them in the corresponding entry of the output root (TFile [2])
63 files. After processing the events, the program exports the obtained topology information at the
64 sample level to the output plain text, tex source, and pdf files.

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

73 The program has a history of more than ten years. It has already gone through a series of
74 major upgrades. Prior to its development, analysts usually wrote some private codes to match
75 few signals and/or backgrounds for their own studies. The limited functions of these codes
76 do not satisfy the increasing demand for topology analysis. This motivates us to develop a
77 generic, powerful, and easy-to-use program. At first, the program was developed for the BESIII
78 experiment, an experiment in the τ -Charm energy region with abundant research topics under
79 study [3, 4]. Later, it was extended substantially for the Belle II experiment, which is primarily
80 dedicated to search for physics beyond the Standard Model in the flavor sector and has already
81 started data taking in the recent three years [5]. Besides, the program has also been tried and
82 used in the Belle experiment, the predecessor of the Belle II experiment, where some physics
83 studies are still ongoing [6]. Not relying on any specific software frameworks, the program now
84 applies to many high energy physics experiments.

85 This user guide gives a detailed description of TopoAna. It proceeds as follows: Section 2
86 introduces the basics of the program; Sections 3 and 4 expatiate the two categories of function-
87 alities of the program — component analysis and signal identification, respectively; Sections 5
88 and 6 present some common settings and auxiliary facilities for the executing of the program,
89 respectively; Section 7 summarizes the user guide. It is worth mentioning here that, aside from
90 the detailed description in the user guide, an essential description of the program can be found in
91 the file “paper_draft_v*.pdf” under the directory “share” of the package.

92 **2. Basics of the program**

93 This section introduces the basics of the program, including the package, input, algorithm,
94 execution, performance, output, and validation of the program. The package implements the

95 program via a C++ class called “topoana” and a main function invoking the class. Compiling
96 the package creates the executable file of the program, that is, “topoana.exe”. To execute the
97 program, we have to first obtain the input data of the program, namely the raw topology truth
98 information of the inclusive MC samples, with some interfaces to the program in the software
99 systems of the corresponding experiments. Normally, the input data contain all the topology
100 information of the samples. With the data, all kinds of the topology analysis presented in the
101 user guide can be performed.

102 To carry out the topology analysis desired in our work, we have to provide some necessary in-
103 put, functionality, and output information to the program. The information is required to be filled
104 in the setting items designed and implemented in the program, and the items have to be put in a
105 plain text file named with a suffix “.card”. With the card file, one can execute the program with
106 the command line: “topoana.exe cardFileName”, where the argument “cardFileName” is option-
107 al and its default value is “topoana.card”. After the execution of the program, we can examine
108 the results of topology analysis in the output files and use them to analyze other experimental
109 quantities. The results help us gain a better understanding of the signals and backgrounds and
110 are conducive to carrying our work forward. Besides the package, input, execution, and output
111 of the program mentioned above, the algorithm, performance, and validation of the program will
112 also be discussed in this section, because they are also essential aspects of the program. In the
113 next seven subsections, we will present the package, input, algorithm, execution, performance,
114 output, and validation of the program in detail, with each part in one subsection.

115 2.1. Package of the program

116 The package consists of six directories — “include”, “src”, “bin”, “share”, “examples”, and
117 “utilities” — and five files — “LICENSE”, “README.md”, “Configure”, “Makefile”, and “Set-
118 up”. While the directory “include” only includes one header file “topoana.h”, the directory “src”
119 contains sixty source files “*.cpp” as well as a script file “topoana.C”. At present, only one class,
120 namely “topoana”, is defined in the program for all of its functionalities. The class is declared in
121 “topoana.h”, implemented in “*.cpp” files, and invoked in “topoana.C”.

122 The file “template_topoana.card” under the directory “share” saves all the items which are de-
123 veloped for users to specify information for the execution of the program. One can refer to the file
124 when filling in the cards for their own needs. Some plain text files “pid_3pchrg_txtpnm_txpnm
125 _iccp.dat_*” are also included in the directory “share”. They store the basic information of the
126 particles used in the program. The suffixes of their names indicate the experiments they apply
127 to. One of them will be copied to “pid_3pchrg_txtpnm_txpnm_iccp.dat” when we set up the
128 program. Besides, the directory “share” also contains three LaTeX style files “ geometry.sty”,
129 “ifxetex.sty”, and “makecell.sty”, which are invoked by the program for generating pdf files.
130 The directory “examples” includes plenty of detailed examples. Particularly, all the examples in-
131 volved in this user guide are under its sub-directory “in_the_user_guide”. The directory “utilities”
132 contains some useful bash scripts.

133 The program is released under MIT license [7]. The file “README.md” briefly introduces
134 how to install and use the program. To set up the program, one should first set the package path
135 with the command “./Configure”. Standard outputs of the command are the guidelines for man-
136 ually adding the absolute path of “topoana.exe” to the environment variable “PATH”, in order to
137 execute it without any path. The second step is executing the command “make”. This command
138 compiles the header, source, and script files into the executable file “topoana.exe” under the di-
139 rectory “bin”, according to the rules specified in the “Makefile”. The last step is specifying the
140 experiment name with the command line “./Setup experimentName”. Currently, the supported

experiment names are “BESIII”, “Belle”, and “Belle-II”. Besides, “./Setup Example” is required for the execution of the examples in the user guide.

143 2.2. *Input of the program*

The input of the program is one or more root files including a TTree [8] object which contains raw topology truth information of the inclusive MC samples under study. To be specific, the information 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 [9] 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 input data.

Number of particles	:	63	(1)	
PDG codes of particles	:	300553, -511, 511, -433, 421, 211, 22, -413, 111, 111, 113, 211, -431, 22, -323, 213, -421, -211, 22, 22, 22, 22, 211, -211, 333, 11, -12, 22, -311, -211, 211, 111, 221, 331, 321, -321, 310, 22, 22, 111, 111, 111, 111, 111, 221, 111, 111, 22, 22, 22, 22, 22, 22, 22, 22, 22, 22, 22, 22, 22, 22, 22, 22, 22, 22, 22		
Mother indices of particles	:	-1, 0, 0, 1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 3, 4, 4, 7, 7, 8, 8, 9, 9, 10, 10, 12, 12, 12, 12, 14, 14, 15, 15, 16, 16, 24, 24, 28, 31, 31, 32, 32, 32, 33, 33, 33, 36, 36, 39, 39, 40, 40, 41, 41, 42, 42, 43, 43, 44, 44, 45, 45, 46, 46		

¹⁴⁴ The complete physics process contained in the data is displayed as follows.

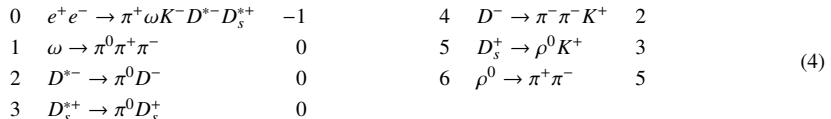
0	$e^+e^- \rightarrow \Upsilon(4S)$	-1	9	$\rho^+ \rightarrow \pi^0\pi^+$	6
1	$\Upsilon(4S) \rightarrow B^0\bar{B}^0$	0	10	$K^{*-} \rightarrow \pi^-\bar{K}^0$	6
2	$B^0 \rightarrow \pi^0\rho^0\pi^+D^{*-}$	1	11	$D_s^- \rightarrow e^-\bar{\nu}_e\phi\gamma$	7
3	$\bar{B}^0 \rightarrow \pi^+D^0D_s^*\gamma$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

145 Here, the decay branches in the process are placed into two blocks in order to make full use of
 146 the page space. In both blocks, the first, second, and third columns are the indices, symbolic
 147 expressions, and mother indices of the decay branches. Notably, all the decay branches of $\pi^0 \rightarrow$
 148 $\gamma\gamma$ are omitted in Eq. (2) in order to make the process look more concise. Since the topology
 149 diagram of such a process looks like a tree, we refer to the complete processes as decay trees.
 150 Obviously, the input data do not show the structure automatically. Thus, we need the program to
 151 do the topology analysis work.

From the first branch in Eq. (2), only one particle $\Upsilon(4S)$ is produced after the e^+e^- annihilation. Thus, $\Upsilon(4S)$ can be referred to as the root particle of the decay tree. Similarly, many other resonances with the quantum numbers $J^{PC} = 1^{--}$, such as J/ψ , can be solely produced at other proper energy points. Besides the cases with only one root particle, the program can deal with the cases with multiple root particles. For example, the program can recognize the following raw topology truth information

Number of particles	:	25	
PDG codes of particles	:	433, -321, 223, 211, -413, 431, 111, 211, -211, 111, -411, 111, 321, 113, 22, 22, 22, 22, 321, -211, -211, 22, 22, 211, -211	(3)
Mother indices of particles	:	-1, -1, -1, -1, -1, 0, 0, 2, 2, 2, 4, 4, 5, 5, 6, 6, 9, 9, 10, 10, 10, 11, 11, 13, 13	

as the following process



152 Here, the particles $\pi^+ \omega K^- D^{*-} D_s^{*+}$ in the first branch arise from hadronization processes, in which
 153 quark pairs produced from initial state particles turn into hadrons. The processes with hadroniza-
 154 tion ignored have a tree structure and thus are easy to resolve. On the other hand, some hadroniza-
 155 tion processes, particularly those in high energy regions, contain complicated loop structures that
 156 are difficult to resolve without sophisticated algorithms. Resolving these intricate hadronization
 157 processes is not involved in the program at present.

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 these 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 [10] object as a scalar integer, while the PDG codes of particles, as well as the mother indices of particles, can be stored in a TBranch object as an array of integers, in a TBranch object as a vector of integers, or in a group of TBranch objects 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 TBranch objects are not recommended to store arrays and vectors in order to use other tools such as NumPy [11] and pandas [12]. In such a situation, we have to store the number of particles in a TBranch object 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 TBranch objects 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).

176 It is easy to get the input of the program within the software framework of high energy
177 physics experiments. To facilitate its use, we have developed the interfaces of the program to
178 the software systems of the BESIII, Belle, and Belle II experiments. Similar interfaces for other

179 experiments can also be implemented with ease. Beyond the scope of the user guide, we will not
180 discuss the details of the interfaces here.

181 *2.3. Algorithm of the program*

The program resolves physics processes from the input data introduced above. Considering the diversity of the data, the program first sorts them before translating them into physics processes. Here, the diversity means that the data representing a process may have multiple permutations. For example, the data for the decay $\rho^0 \rightarrow \pi^+\pi^-$ have the following two permutations.

Number of particles	:	3
PDG codes of particles	:	<u>113, 211, -211</u> or <u>113, -211, 211</u>
Mother indices of particles	:	-1, 0, 0

182 A decay tree can consist of many decay branches. As a consequence, the diversity issue is
183 complex. To avoid the different permutations of one group of data are identified as different pro-
184 cesses, the program first sorts the input data to adjust all the possible permutations to a unique
185 order, according to the PDG codes and electronic charges of the involved particles, and the num-
186 bers of their daughter particles in the case of identical particles present in the same decay branch.
187 For example, the two permutations above will be finally sorted into the first permutation (113,
188 211, -211) in the program. The sorting algorithm is implemented in the source file “sortPs.cpp”,
189 where some other settings are also involved. One can see the reference file “sortPs.cpp_core” for
190 the core of the sorting algorithm. After the sorting, the program can get the decay tree from the
191 sorted data into a vector of the type “vector< list<int> >” with the function implemented in the
192 source file “getDcyTr.cpp”.

193 As mentioned in the previous section, the program has two categories of functionalities: sig-
194 nal identification and component analysis. In this subsection, we introduce the basic algorithms
195 for signal identification and component analysis by taking the cases of decay trees as examples.
196 Figures 2 and 3 show the flow charts of these algorithms in detail. Dozens of lines of code, in-
197 cluding some using the ROOT classes TChain [13], TFile [2], and TTree [8], are involved in the
198 charts in order to express the algorithms explicitly. The flow chart of the signal identification for
199 decay trees is depicted in Fig. 2. Firstly, the program reads in the signal decay trees specified in
200 the user card file. Then, for each entry of the input root file, the program obtains the decay tree
201 from the sorted input data, matches the decay tree to the signal decay trees, records the index of
202 the matched signal decay tree, and increases the number of the matched signal decay tree. At
203 last, the program outputs the statistics of the signal decay trees.

204 The flow chart of the component analysis over decay trees is illustrated in Fig. 3. Despite
205 the similarity in their frameworks, the flow chart has significant differences from that of the
206 signal identification for decay trees in Fig. 2. In the signal identification algorithm, the signal
207 decay trees to be identified are specified beforehand in the user card file. On the contrary, in
208 the component analysis algorithm, the program has to classify decay trees by itself from scratch.
209 In the signal identification algorithm, the decay trees are matched by directly comparing the
210 vectors storing them. Since the number of specified signal decay trees is fixed and usually small,
211 the processing rate of the program is high and usually in constant. However, in the component
212 analysis algorithm, the number of decay tree types found in a sample can be quite large and tends
213 to grow with the number of processed entries. On this occasion, if we still match the decay trees
214 by comparing the vectors storing them, the processing rate of the program will decrease with
215 the increase of the number of processed entries. To improve the processing rate, the unordered
216 map [14], a kind of container template introduced since the C++ 11 standard, is employed for the

217 fast matching of decay trees. Internally, the elements in the unordered maps are organized into
 218 buckets depending on their hash values, to allow for fast access to individual elements directly by
 219 their key values with a constant average time complexity [14]. This constant feature in average
 220 time complexity will be examined in Section 2.5.

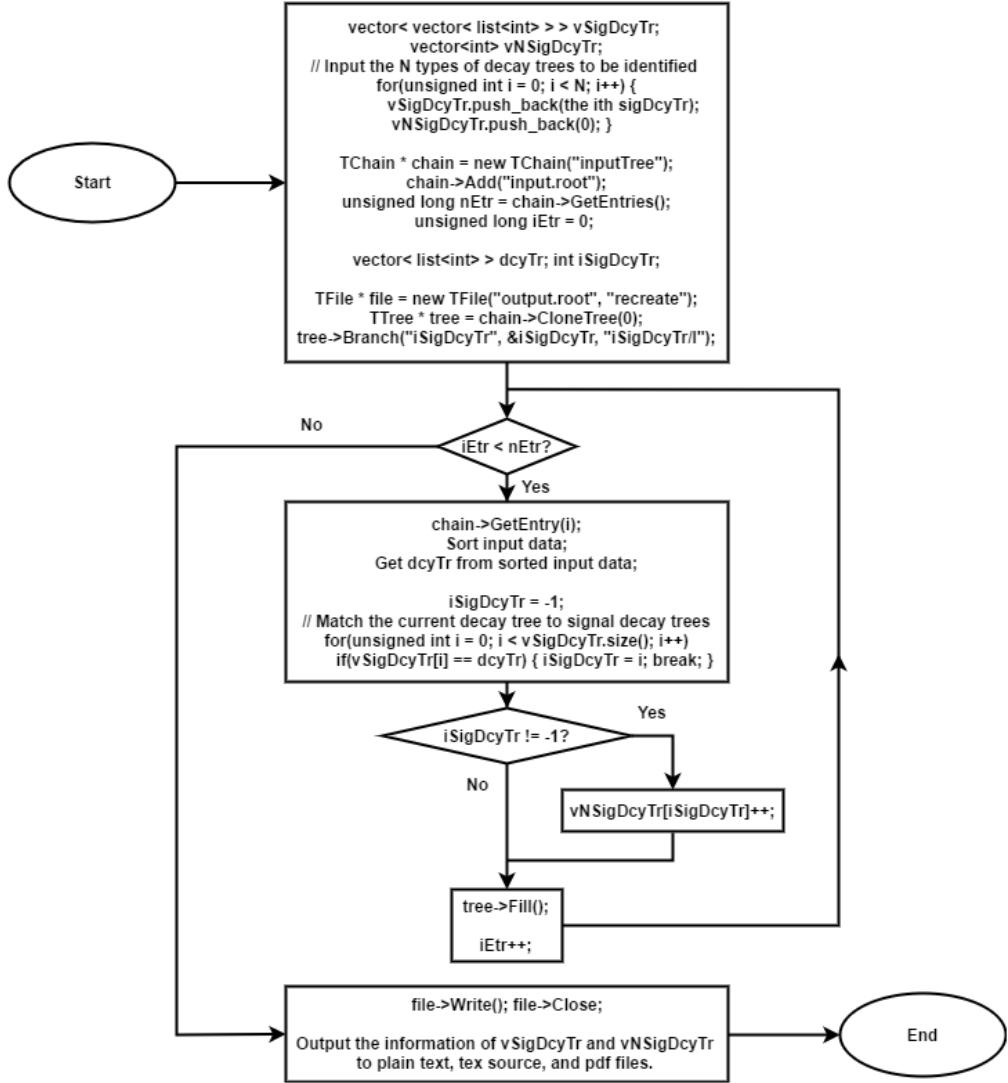


Figure 2: Basic flow chart of the signal identification for decay trees. The vectors “vSigDcyTr” and “vNSigDcyTr” are used to store the signal decay trees specified in the user card file and the numbers of these decay trees found in the input root file, respectively. The TBranch “iSigDcyTr” in the output root file is used to record the index of the signal decay tree involved in each entry of the input root file.

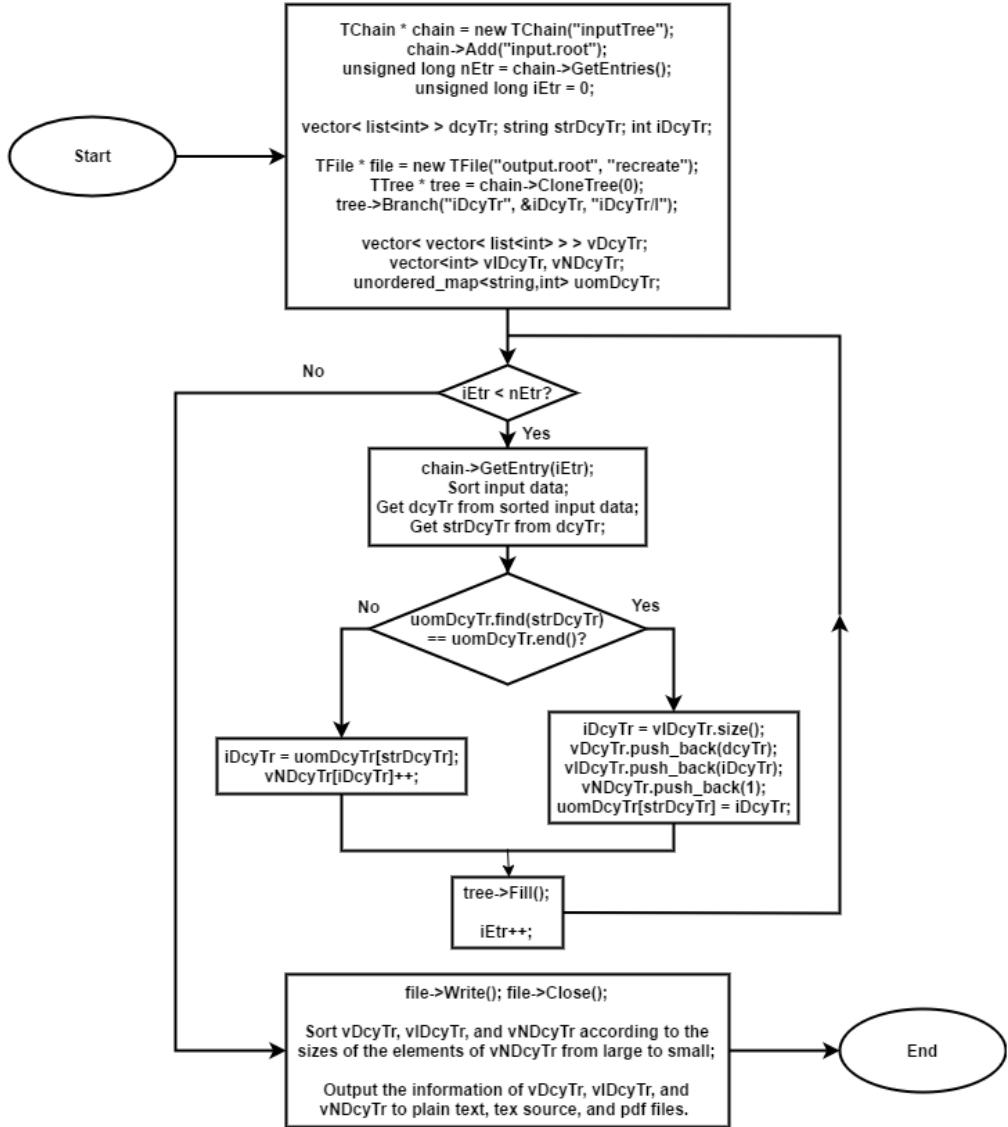


Figure 3: Basic flow chart of the component analysis over decay trees. The TBranch “iDcyTr” in the output root file is used to record the index of the decay tree involved in each entry of the input root file. The vectors “vDcyTr”, “vIDcyTr”, and “vNDcyTr” are used to store the decay trees found in the input root file, their individual indices, and their individual numbers, respectively. In addition, the unordered_map “uomDcyTr” is used for the fast matching of decay trees. Its key and value are the string “strDcyTr” and the index “iDcyTr”, respectively. Here, the string “strDcyTr” is constructed from the vector “dcyTr”; there is a one-to-one correspondence between them.

221 2.4. Execution of the program

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

228 An example of the card file containing the essential items is shown as follows.
229

```
230
231 # The following six items set the input of the program.
232
233 % Names of input root files
234 {
235     ./input/jpsi_1.root
236     ./input/jpsi_2.root
237 }
238
239 % TTree name
240 {
241     evt
242 }
243
244 % Storage type of input raw topology truth information (Four options: AOI, VOI, MSI, and MSD. Default: AOI)
245 {
246     AOI
247 }
248
249 % TBranch name of the number of particles (Default: nMCGen)
250 {
251     Nmcps
252 }
253
254 % TBranch name of the PDG codes of particles (Default: MCGenPDG)
255 {
256     Pid
257 }
258
259 % TBranch name of the mother indices of particles (Default: MCGenMothIndex)
260 {
261     Midx
262 }
263
264 # The following item sets the basic functionality of the program.
265
266 % Component analysis — decay trees
267 {
268     Y
269 }
270
271 # The following item sets the output of the program.
272
273 % Common name of output files (Default: Name of the card file)
274 {
275     jpsi_ta
276 }
```

277
278 In the card file, “#”, “%”, and the pair of “{” and “}”, are used for commenting, prompting,
279 and grouping, respectively. The first six, seventh, and last items are set for the input, basic
280 functionality, and output of the program, respectively.

281 The first item sets the names of the input root files. The names ought to be input one per
282 line without tailing characters, such as comma, semicolon, and period. In the names, both the
283 absolute and relative paths are allowed and wildcards “[]?*” are supported, just like those in the
284 root file names input to the method Add() of the class TChain [13]. The second item specifies
285 the TTree name. The third item tells the program the storage type of the input raw topology truth
286 information, and the input should be one of the following four acronyms: AOI, VOI, MSI, and
287 MSD, as we introduce in the previous subsection. The following three items set the TBranch
288 names of the three ingredients of the input raw topology truth information. Of the first six items,
289 the former two are indispensable, whereas the latter four can be removed or left empty if the
290 input values are identical to the default values indicated in their prompts. Besides, the latter four
291 items can be moved to the underlying card file, which is developed for frequently used items and
292 will be introduced in Section 6.1, because the input values are usually fixed for a user or a group
293 of users, though they might be different from the default values.

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

299 The last item specifies the common name of the output files. Though in different formats, the
300 files are denominated with the same name for the sake of uniformity. They will be introduced
301 in detail in the next subsection. This item is also optional, with the name of the card file as its
302 default input value. It is a good practice to first denominate the card file with the desired common
303 name of the output files and then remove this item or leave it empty.

304 To provide a complete description, we list and explain all the essential items in the paragraphs
305 above. However, in practical uses, we suggest removing the optional items if the input values
306 are identical to the default ones, or moving them to the underlying card file if the input values
307 are fixed for most of your use cases. In this way, the contents of the card file will become much
308 more concise, making the use of the program easier and quicker. For example, unless otherwise
309 stated, only the following two items are used to set the essential information in Sections 3, 4, and
310 5.

```
311 % Names of input root files
312 {
313     ./input/mixed_1.root
314     ./input/mixed_2.root
315 }
316
317 % TTree name
318 {
319     evt
320 }
321
```

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

325 During the execution of the program, some standard output and error messages are printed to

327 the screen to provide some information on the input, progress, and output of the program, as well
 328 as the possible problems and proposed solutions to them. The standard output messages include
 329 the following four parts: (1) the values of the items with active inputs; (2) the total number of
 330 entries contained in the input root files and the progress of the program to process these entries;
 331 (3) the information output by the pdflatex command when it compiles the tex source file to
 332 get the pdf file; (4) and the hints on the output of the program. The standard error messages
 333 are prompted with “Error:” and “Infor:” in order to differentiate themselves from the standard
 334 output messages. The messages started with “Error:” point out the problems encountered by the
 335 program directly, while those started with “Infor:” give more information on the problems as
 336 well as some guidelines on the solutions to the problems.

337 2.5. Performance of the program

338 Besides the performance of the used computing systems, the processing rate of the program is
 339 largely related to the characteristics of the samples, particularly the average number of generated
 340 particles in each event. Figure 4 shows the performance study of the program with the J/ψ
 341 sample used in the example of this section as well as the $\tau^+\tau^-$, $d\bar{d}$, $u\bar{u}$, $s\bar{s}$, $c\bar{c}$, B^+B^- , and $B^0\bar{B}^0$
 342 samples generated at the peak energy of the $\Upsilon(4S)$ resonance. From the left plot in the figure,
 343 for all the samples, the number of elapsed seconds grows linearly with the number of processed
 344 entries. This linear pattern is a nice feature. It guarantees the program has a high rate even in the
 345 case of processing huge samples. For example, the program can process one hundred thousand
 346 J/ψ events within five seconds. Here, we note that the linear pattern is the result of fast searches
 347 with unordered maps [14], as we discuss in Section 2.3. On the other hand, the processing rate
 348 of the program varies with the processed samples. The right plot in Fig. 4 shows the relationship
 349 between the total number of elapsed seconds over the whole sample and the average number
 350 of generated particles in an event. Clearly, a linear pattern is also observed in the plot. To be
 351 specific, with the average number of generated particles in an event increasing by one, the total
 352 number of elapsed seconds over the whole sample increases by about 0.56.

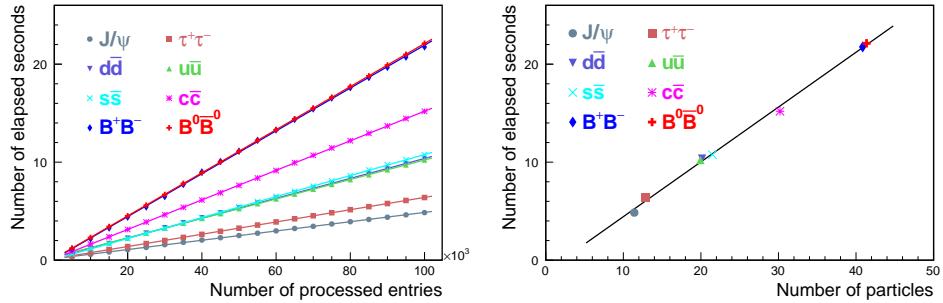


Figure 4: Performance study of the program with the J/ψ sample as well as the $\tau^+\tau^-$, $d\bar{d}$, $u\bar{u}$, $s\bar{s}$, $c\bar{c}$, B^+B^- , and $B^0\bar{B}^0$ samples generated at the peak energy of the $\Upsilon(4S)$ resonance. The left plot demonstrates the changing trends of the number of elapsed seconds with respect to the number of processed entries. The right plot illustrates the relationship between the total number of elapsed seconds over the whole sample and the average number of generated particles in an event. In both plots, the dots show the timing data from the standard output of the program, and the lines display the results of fitting linear functions to the data.

353 2.6. *Output of the program*

354 The program gains the topology information from input data and saves it to output files. As
355 mentioned in Section 1, the information includes the types of physics processes and the number
356 of processes in each type, involved both in entire samples and in individual events. We refer to
357 the information at the sample level as topology maps. In the topology maps, we assign an integer
358 to each type of physics processes as its index. We term the indices of processes as well as the
359 numbers of processes involved in each type in the individual events as topology tags.

360 The program outputs topology maps to three different files: one plain text file, one tex source
361 file, and one pdf file, with the same name specified in the card file. For instance, the three files
362 are “jpsi_ta.txt”, “jpsi_ta.tex”, and “jpsi_ta.pdf” in the example. Although in different formats,
363 the three files have the same information. The pdf file is the easiest to read. It is converted from
364 the tex source file with the command pdflatex. The tex source file is convenient to us if we want
365 to change the style of the pdf file to our taste and when we need to copy and paste (parts of) the
366 topology maps to our slides, papers, and so on. For example, all of the tables displaying topology
367 maps in this user guide are taken from associated tex source files. The plain text file has its own
368 advantage, because the topology maps in it can be checked with text processing commands as
369 well as text editors, and can be used on some occasions as input to the functionality items (see
370 Sections 3 and 4 for details) of another card file.

371 In addition to the three files for topology maps, one or more root files are output to save
372 topology tags. The root files only include one TTree object, which is entirely the same as that in
373 the input root files, except for the topology tags inserted in all of its entries. The number of root
374 files depends on the size of output data. The program switches to one new root file whenever the
375 size of the TTree object in memory exceeds 3 GB. In the case of the size less than 3 GB, only
376 one root file is output. While the sole or first root file has the same name as the three files above,
377 more possible root files are denominated with the suffix “_n” (n=1, 2, 3, and so on) appended to
378 the name. In the example, the first root file is “jpsi_ta.root”, and more possible root files would
379 be “jpsi_ta_1.root”, “jpsi_ta_2.root”, “jpsi_ta_3.root”, and so on.

380 In the example of the previous subsection, the program conducts its basic functionality, namely
381 the component analysis over decay trees. From the 100000 events of the input sample, the
382 program recognizes 17424 decay trees and outputs all of them to the plain text, tex source, and
383 pdf files. Table 1 only shows the top ten decay trees and their respective final states listed in the
384 output pdf file. With the help of the symbolic expressions, the components of the sample are
385 clearly displayed in the table, which brings great convenience to us in examining the signals and
386 backgrounds involved in the sample. In the table, “rowNo”, “iDcyTr”, “nEtr”, and “nCEtr” are
387 abbreviations for the row number, index of decay tree, number of entries of decay tree, and num-
388 ber of the cumulative entries from the first to the current decay trees, respectively. The values of
389 “iDcyTr” are assigned from small to large in the program but listed according to the values of
390 “nEtr” from large to small in the table. This is the reason why they are not in natural order like
391 the values of “rowNo”. Since J/ψ is the only root particle for the J/ψ sample, the production
392 branch $e^+e^- \rightarrow J/\psi$ is omitted to save page space. Similar rules also apply to other samples with
393 only one root particle. Considering π^0 has a very large production rate and approximatively 99%
394 of it decays to $\gamma\gamma$, the program is designed to discard the decay $\pi^0 \rightarrow \gamma\gamma$ by default at the early
395 phase of processing the input data (see Section 5.1.2 for the setting item to alter the behavior).
396 As a result, $\pi^0 \rightarrow \gamma\gamma$ does not show itself in the table. Besides, the superscripts “*f*” and “*F*” in
397 γ^f and γ^F indicate the final state radiation effect (see Section 5.1.3 for their difference).

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	$J/\psi \rightarrow \pi^0 \pi^+ \pi^+ \pi^- \pi^-$	$\pi^0 \pi^+ \pi^+ \pi^- \pi^-$	0	2850	12632
4	$J/\psi \rightarrow \pi^0 \pi^+ \pi^+ \pi^- \pi^- \pi^-$	$\pi^0 \pi^+ \pi^+ \pi^- \pi^- \pi^-$	2	1895	14527
5	$J/\psi \rightarrow \pi^0 \pi^+ \pi^- K^+ K^-$	$\pi^0 \pi^+ \pi^- K^+ K^-$	20	1698	16225
6	$J/\psi \rightarrow \rho^+ \rho^- \omega, \rho^+ \rightarrow \pi^0 \pi^+, \rho^- \rightarrow \pi^0 \pi^-, \omega \rightarrow \pi^0 \pi^+ \pi^-$	$\pi^0 \pi^0 \pi^+ \pi^+ \pi^- \pi^-$	19	1453	17678
7	$J/\psi \rightarrow e^+ e^- \gamma^f$	$e^+ e^- \gamma^f$	70	1222	18900
8	$J/\psi \rightarrow \pi^0 \pi^0 \pi^+ \pi^+ \pi^- \pi^-$	$\pi^0 \pi^0 \pi^+ \pi^+ \pi^- \pi^-$	127	1161	20061
9	$J/\psi \rightarrow \pi^0 \pi^+ \pi^+ \pi^+ \pi^- \pi^- \pi^- \pi^-$	$\pi^0 \pi^+ \pi^+ \pi^+ \pi^- \pi^- \pi^- \pi^-$	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

398 In the table, “iDcyTr” is the topology tag for decay trees. Thus, it is also saved in the TTree
399 objects of the output root file, together with other quantities for physics analysis. Therefore, it
400 can be used to pick out the entries of specific decay trees and then examine the distributions of
401 the other quantities over the decay trees. In the example, besides the raw topology truth informa-
402 tion, only a random variable following the standardized normal distribution, namely X, is stored
403 in the input root files and thus copied by default to the output root file. Though not a genuine
404 variable for physics analysis, X is quite good to illustrate the usage of the topology tag. Figure 5
405 shows the distribution of X accumulated over the top ten decay trees. The figure is drawn with
406 the root script

407
408 examples/in_the_user_guide/ex_for_tb_01/draw_X/v2/draw_X.C,
409

410 where, for example, a statement equivalent to

411
412 chain->Draw(“X >>h0”, “iDcyTr==6”)

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

420 2.7. Validation of the program

421 The decay trees displayed in Table 1 are relatively simple, and we can check their correctness
422 by examining the input data directly. To validate the program generally, we need to do input and
423 output checks, where some arbitrary physics processes are generated as the input of the program.
424 The output has to be consistent with the input; otherwise, there must be some bugs in the program
425 and we have to fix them. A large number of such checks have been performed in the develop-
426 ment and application of the program, and some of them can be found under the sub-directory
427 “examples/validation” of the package. These checks are divided into two groups: standalone and
428 combined. In the standalone checks, forty exclusive J/ψ and $\Upsilon(4S)$ decays modeled with the

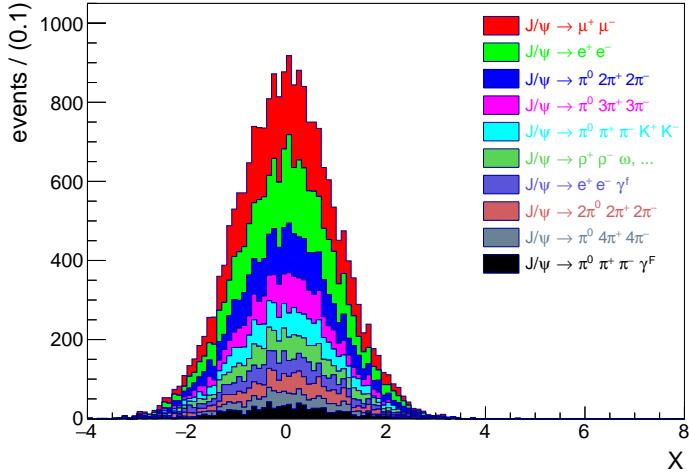


Figure 5: Distribution of X accumulated over the top ten decay trees. In the legend entry “ $J/\psi \rightarrow \rho^+ \rho^- \omega, \dots$ ”, the dots “ \dots ” represent the secondary decay branches: $\rho^+ \rightarrow \pi^0 \pi^+$, $\rho^- \rightarrow \pi^0 \pi^-$, $\omega \rightarrow \pi^0 \pi^+ \pi^-$.

429 EvtGen [15] generator are used to test the functionality of resolving decay trees. In the combined
 430 checks, randomly combined samples of these exclusive decays are used for verifying the
 431 functionalities of counting and tagging decay trees. The output agrees with the input in all the
 432 checks, which indicates the correctness of the program.

433 3. Component analysis

434 Component analysis is the primary functionality of the program. It is developed mainly for
 435 the background analysis involved in our physics studies. We perform it over decay trees in the
 436 previous example. Also, it can be carried out as follows: over decay initial-final states; with
 437 specified particles to check their decay branches, production branches, mothers, cascade decay
 438 branches, and decay final states; with specified inclusive decay branches to examine their exclu-
 439 sive components; and with specified intermediate-resonance-allowed (IRA) decay branches to
 440 investigate their inner structures. This section introduces the nine (five for specified particles)
 441 kinds of component analysis, with each in a subsection. For each kind of component analysis,
 442 one item is designed and implemented in the program to set related parameters. In each subsec-
 443 tion, we take an example to demonstrate the corresponding setting item and show the resulting
 444 topology map. For easy exposition, all of the essential topology tags involved in the component
 445 analysis functionalities are presented in another separate subsection, namely the last subsection.

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

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

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

470 3.1. Decay trees

471 Component analysis over decay trees is the basic kind of topology analysis. It is quite useful
 472 to study the backgrounds involved in our research works where the signals are the complete decay
 473 trees fully reconstructed from final state particles. It has already been widely performed in the
 474 BESIII experiment, as illustrated in the previous section with the J/ψ example. This subsection
 475 introduces it further with the available optional settings using the $\Upsilon(4S)$ sample. The following
 476 example shows the associated item with the maximum number of output components set to five.
 477 In the item, a third parameter is also filled and set to “Y”. With the setting, the decay final states
 478 in the output pdf file are put under their respective decay trees, rather than in a column next to
 479 that for decay trees. It is recommended to use this optional parameter in cases there are too many
 480 (about ten or more) particles in some final states. Here, we note that the symbol “–” can be used
 481 as a placeholder for the maximum number of output components, if only the third parameter is
 482 desired.

```

483
484
485        % Component analysis — decay trees
486        {
487            Y   5   Y
488        }
  
```

489 Component analysis over decay trees is one kind of the most time-consuming topology anal-
 490 ysis tasks. To check further the efficiency of the program, the progress of running this example,
 491 in addition to the example in Section 2.4, is illustrated in the plots of Fig. 4 as well. In these plots,
 492 the timing data from this example are marked with the legend entry “ $B^0\bar{B}^0$ ”. Since the decay of
 493 the $\Upsilon(4S)$ resonance is more complex than that of the J/ψ resonance, it takes more than twenty
 494 seconds for the program to process one hundred thousand events in this example. Nonetheless,
 495 the program still has a high processing rate.

496 Table 2 shows the decay trees. In the table, while the first five decay trees are listed exclu-
 497 sively in the main part, the rest decay trees are only summarized inclusively at the bottom row.
 498 Here, we note that the events are not densely populated over the first five decay trees because the

500 inclusive $\Upsilon(4S)$ sample used here is not selected beforehand with any requirements. In the sym-
 501 bolic expressions of decay initial-final states, the dashed right arrow (\dashrightarrow) instead of the plain
 502 right arrow (\rightarrow) is used, in order to reflect that the initial states do not necessarily decay to the
 503 final states in a direct way. Similarly, it is also used in the symbolic expressions of IRA decay
 504 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	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow e^+ \nu_e D^{*-} \gamma^F, \bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}, D^{*-} \rightarrow \pi^- \bar{D}^0,$ $D^{*+} \rightarrow \pi^+ D^0, \bar{D}^0 \rightarrow \pi^0 \pi^- K^+, D^0 \rightarrow \pi^0 \pi^+ K^-$ $(\Upsilon(4S) \dashrightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu \pi^0 \pi^0 \pi^+ \pi^+ \pi^- \pi^- K^+ K^- \gamma^F)$	20870	3	3
2	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow \pi^0 \pi^+ \pi^+ \rho^- D^-, \bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}, \rho^- \rightarrow \pi^0 \pi^-,$ $D^- \rightarrow \pi^- \pi^- K^+, D^{*+} \rightarrow \pi^+ D^0, D^0 \rightarrow K_L^0 \pi^+ \pi^-$ $(\Upsilon(4S) \dashrightarrow \mu^- \bar{\nu}_\mu \pi^0 \pi^0 K_L^0 \pi^+ \pi^+ \pi^+ \pi^- \pi^- \pi^- K^+)$	5295	2	5
3	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow \mu^+ \nu_\mu D^{*-}, \bar{B}^0 \rightarrow e^- \bar{\nu}_e D^+, D^{*-} \rightarrow \pi^- \bar{D}^0,$ $D^+ \rightarrow e^+ \nu_e \bar{K}^*, \bar{D}^0 \rightarrow \pi^0 \pi^+ \pi^- K_S^0, \bar{K}^* \rightarrow \pi^0 \bar{K}^0, K_S^0 \rightarrow \pi^+ \pi^-, \bar{K}^0 \rightarrow K_L^0$ $(\Upsilon(4S) \dashrightarrow e^+ e^- \nu_e \bar{\nu}_e \mu^+ \nu_\mu \pi^0 \pi^0 K_L^0 \pi^+ \pi^+ \pi^- \pi^- \pi^-)$	11954	2	7
4	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow e^+ \nu_e D^{*-}, \bar{B}^0 \rightarrow \pi^0 \pi^- \omega D^+, D^{*-} \rightarrow \pi^- \bar{D}^0,$ $\omega \rightarrow \pi^0 \pi^+ \pi^-, D^+ \rightarrow e^+ \nu_e \pi^+ K^-, \bar{D}^0 \rightarrow \pi^0 \pi^- K^+$ $(\Upsilon(4S) \dashrightarrow e^+ e^+ \nu_e \pi^0 \pi^0 \pi^0 \pi^+ \pi^+ \pi^- \pi^- \pi^- K^+ K^-)$	14345	2	9
5	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow \mu^+ \nu_\mu D^{*-}, \bar{B}^0 \rightarrow e^- \bar{\nu}_e D^{*+} \gamma^F, D^{*-} \rightarrow \pi^- \bar{D}^0,$ $D^{*+} \rightarrow \pi^0 D^+, \bar{D}^0 \rightarrow \pi^- K^+, D^+ \rightarrow e^+ \nu_e \bar{K}^*, \bar{K}^* \rightarrow \pi^+ K^-$ $(\Upsilon(4S) \dashrightarrow e^+ e^- \nu_e \bar{\nu}_e \mu^+ \nu_\mu \pi^0 \pi^+ \pi^+ \pi^- \pi^- K^+ K^- \gamma^F)$	15332	2	11
rest	$\Upsilon(4S) \rightarrow \text{others (99980 in total)}$ $(\Upsilon(4S) \dashrightarrow \text{corresponding to others})$	—	99989	100000

505 3.2. Decay initial-final states

Table 3: Decay initial-final states.

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

506 On some occasions, we need to investigate the decay initial-final states of backgrounds for
 507 some sophisticated physics analyses. Particularly, it is necessary to differentiate the following
 508 two fundamental types of backgrounds: the one with the same initial-final states as the signal,
 509 and the other with different initial-final states from the signal. While the latter type of back-
 510 grounds needs to be suppressed as much as possible, the former type usually needs to be kept to
 511 study more physical effects, for example, the interference effect. Besides, examining the decay
 512 initial-final states of backgrounds sheds light on the misjudgment of final state particles at the
 513 level of signal candidates. Below is an example demonstrating the related item with the maxi-
 514 mum number of output components set to five.

515 % Component analysis — decay initial-final states

```

517   {
518     Y 5
519   }
520

```

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

523 3.3. Decay branches of particles

524 The invariant mass constraint is one of the most frequently used event selection requirements
 525 in high energy physics experiments. With the requirement applied to certain particle, the main
 526 backgrounds (especially the peaking ones) to its signal decay mode are very likely to be its other
 527 decay modes. In this case, it is significant to examine the decay branches of the particle. The
 528 following example shows the associated item with the two particles D^{*+} and J/ψ set as research
 529 objects. In the item, each row holds the information of a specified particle, and the first, sec-
 530 ond and third columns are the textual expressions, aliases, and maximum numbers of output
 531 components, respectively. As we introduce at the beginning part of this section, the aliases and
 532 maximum numbers of output components are both optional. Here, we note that the symbol “–”
 533 can be used as a placeholder for an unassigned alias, if only the maximum number of output
 534 components is desired.

```

535 % Component analysis — decay branches of particles
536 {
537   D*+  Dsp  5
538   J/psi Jpsi 5
539 }
540

```

541 Table 4 shows the decay branches of D^{*+} . From the table, only four decay branches of D^{*+} are
 542 found in the input inclusive MC sample. Since there is likely one or more cases of D^{*+} decays in
 543 one input entry, “nCase” and “nCCase”, instead of “nEtr” and “nCEtr”, are used in the table in
 544 order to accurately indicate what we are counting are the numbers of D^{*+} decays, rather than the
 545 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

547 3.4. Production branches of particles

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

```

551 % Component analysis — production branches of particles
552 {
553   D*+  Dsp  5
554   J/psi Jpsi 5
555 }
556

```

557 558 The production branches of D^{*+} are displayed in Table 5. In the production branches, D^{*+} is

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

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	$\bar{B}^0 \rightarrow \pi^0 \pi^+ \pi^- \pi^- D^{*+}$	40	1429	11783
rest	others (3272 in total)	—	34103	45886

562 3.5. Mothers of particles

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

```
566
567 % Component analysis — mothers of particles
568 {
569   D*+  Dsp  5
570   J/psi Jpsi 5
571 }
```

572 Table 6 shows the mothers of D^{*+} . Notably, the PDG codes of the mother particles, instead of
 573 additional indices, are listed in the table, since they are sufficient to tag the mother particles.
 574 From the table, six sources of D^{*+} are found in the input sample and the dominant one is the \bar{B}^0
 575 decay.
 576

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

577 3.6. Cascade decay branches of particles

578 Sometimes, the invariant mass constraint is applied to certain particle and the signal pro-
 579 cess is its cascade decay branch. In this case, it is necessary to investigate the cascade decay
 580 branches of the particle, rather than its first decay branches, so as to analyze the background-
 581 s effectively. Below is an example demonstrating the related item by taking the two particles
 582 B^0 and D^0 as objects of study. While the first three columns of the input to this item have the
 583 same meanings as those to the three items above, the additional fourth column sets the maximum
 584 hierarchy of decay branches to be examined. Here, the hierarchy reflects the rank of a decay
 585 branch in a cascade decay branch of one specific particle. For instance, in the following cascade
 586 decay branch of B^0 : $B^0 \rightarrow \pi^0 \pi^0 \rho^0 \pi^+ D^{*-}$, $\rho^0 \rightarrow \pi^+ \pi^-$, $D^{*-} \rightarrow \pi^- \bar{D}^0$, $\bar{D}^0 \rightarrow \eta \eta'$, $\eta \rightarrow \pi^0 \pi^0 \pi^0$,

587 $\eta' \rightarrow \pi^0 \pi^0 \eta$, $\eta \rightarrow \gamma \gamma$, the hierarchies of the seven individual decay branches are 1, 2, 2, 3, 4, 4,
 588 and 5, respectively. In the example, the maximum hierarchy of decay branches is set to two for
 589 both B^0 and D^0 , and hence only the first two hierarchies of branches in their cascade decays will
 590 be investigated. Without such settings, all the branches in their cascade decays will be examined.
 591

```

    592    % Component analysis — cascade decay branches of particles
    593    {
    594     B0    B0    5    2
    595     D0    D0    5    2
    596   }
    597
  
```

598 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

599 3.7. Decay final states of particles

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

600 When the invariant mass constraint is applied to certain particle reconstructed directly from
 601 a specific final state, it is significant to examine the decay final states of the particle, rather than
 602 its first or cascade decay branches, in order to study the backgrounds effectively. The following
 603 example shows the associated item also with the two particles B^0 and D^0 set as research objects.
 604 The format of the input to the item is the same as that to the above item, but the fourth parameters
 605 here are designed to restrict the numbers of final state particles. Without the fourth parameters,
 606 all the decay final states of the specified particles will be investigated. In the example, the pa-
 607 rameters are set to three for both B^0 and D^0 , and thus only the three-body decay final states of
 608 them will be examined.

```

    609    % Component analysis — decay final states of particles
    610    {
    611     B0    B0    5    3
    612     D0    D0    5    3
    613   }
    614
  
```

615

616 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
617 D^0 .
618

619 3.8. Inclusive decay branches

620 In a few physics studies, we take inclusive decay branches as signals. In such cases, it is es-
621 sential to have a basic knowledge of the exclusive components of these inclusive decay branches.
622 Below is an example demonstrating the related item by investigating the exclusive components
623 of the two inclusive decay branches $\bar{B}^0 \rightarrow D^{*+} + \text{anything}$ and $B^0 \rightarrow K_S^0 + \text{anything}$. In the
624 item, each row holds the information of an inclusive decay branch, and the first, second, and
625 third columns separated with the symbol “&” are the textual expressions, aliases, and maximum
626 numbers of output components, respectively. As we introduce at the beginning part of this sec-
627 tion, the aliases and maximum numbers of output components are both optional. Here, we note
628 that the symbol “–” can be used as a placeholder for an unassigned alias, if only the maximum
629 number of output components is desired.
630

```
631 % Component analysis — inclusive decay branches
632 {
633   B0 --> D*+ & B2Dsp & 5
634   B0 --> K_S0 & B2Ks & 5
635 }
```

636
637 The exclusive components of $B^0 \rightarrow K_S^0 + \text{anything}$ are displayed in Table 9. From the table,
638 ten exclusive components of the inclusive decay branch are found in the input sample, and the
639 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

640 3.9. Intermediate-resonance-allowed decay branches

641 In many research works, we take multi-body decay branches as signals. On such occasions,
642 it is fundamental to investigate the intermediate resonances involved in these decay branches.
643 In other words, we need to examine the exclusive components of these IRA decay branches.
644 The following example shows the associated item with the two IRA decay branches $D^{*+} \dashrightarrow$
645 $\pi^0\pi^+\pi^-$ and $J/\psi \dashrightarrow \pi^0\pi^+\pi^-$ set as objects of study. Since IRA decay branches look like
646 inclusive decay branches, the format of the input to the item for IRA decay branches is identical
647 to that for inclusive decay branches, which is introduced in the previous subsection.
648

```
649 % Component analysis — intermediate-resonance-allowed decay branches
650 {
651   D*+ --> K- pi+ pi+ pi0 & Dsp2K3Pi & 5
652   J/psi --> pi+ pi- pi0 & Jpsi23Pi & 5
653 }
```

654
655 Table 10 shows the exclusive components of $D^{*+} \dashrightarrow \pi^0\pi^+\pi^-$. From the table, two interme-

656 diate particles D^0 and D^+ are found in the IRA decay branch, and they decay to $\pi^0\pi^+K^-$ and
 657 $\pi^+\pi^+K^-$, 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	$D^{*+} \rightarrow \pi^0D^+, D^+ \rightarrow \pi^+\pi^+K^-$	1	1102	4971

658 3.10. Essential topology tags

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

Component type	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 s
	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 s
	iDcyBrIRADcyBr_i_j	index of decay branch of the j th IRA decay branch _i

659 Table 11 lists and interprets all of the essential topology tags involved in the component
 660 analysis functionalities. The topology tag for the component analysis over decay initial-final
 661 states is iDcyIFsts. It has a similar interpretation as iDcyTr and is shown in the third column
 662 of Table 3. For the latter seven kinds of component analysis, there are two sorts of topology
 663 tags. The first sort, such as nPDcyBr_i, records the number of instances of the ith specified
 664 particle or decay branch found in each event. The second sort, for example, iDcyBrP_i_j, keeps
 665 the associated index of the jth found instance of the ith specified particle or decay branch. The
 666 indices and the decays they stand for can be found in Tables 4 – 10.

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

674 In addition, “j” in “_j” is the default index of the found instance of certain particle or decay
 675 branch in an event, and it ranges from 0 (included) to the sample-level maximum of the number

676 of the particles or decay branches found in each event (excluded). For example, the maximum of
 677 the number of D^{*+} found in each event is two for the whole sample, and thus two topology tags
 678 iDcyBrP_Dsp_0 and iDcyBrP_Dsp_1 are employed to store the indices of D^{*+} decay branches.
 679 These indices range from 0 (included) to the number of the types of D^{*+} decay branches found
 680 in the samples (excluded). In the events with only one D^{*+} , iDcyBrP_Dsp_1 is assigned with
 681 the default value -1; in the events that have no D^{*+} , the default value -1 is assigned to both
 682 iDcyBrP_Dsp_0 and iDcyBrP_Dsp_1. We note that different from all other indices, PDGMoth_i_j
 683 has the default value 0, instead of -1.

684 4. Signal identification

685 Signal identification is the other functionality of the program. Though relatively simple, it
 686 can help us identify the “signals” we desire directly, quickly, and easily. Here, the “signals”
 687 are not confined to the authentic signals in our research works but can be any physics processes
 688 of interests, particularly some important backgrounds we concern. At present, the following
 689 eight kinds of signals can be identified with the program: (1) decay trees, (2) decay initial-final
 690 states, (3) particles, (4) (regular) decay branches, (5) cascade decay branches, (6) inclusive decay
 691 branches, (7) inclusive cascade decay branches, and (8) IRA decay branches. For each kind of
 692 signals, one item is developed to specify related parameters. This section introduces the eight
 693 kinds of signal identification, with each in a subsection. In each subsection, we take an example
 694 to demonstrate the related setting item and show the obtained topology map. For easy exposition,
 695 all of the essential topology tags involved in the signal identification functionalities are presented
 696 in another separate subsection, that is, the last subsection.

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

702 4.1. Decay trees

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

```

714
715     % Signal identification — decay trees
716     {
717         0 & Upsilon(4S) --> B0 anti-B0 & -1 & 1stDcyTrInTb2 & Y
718         1 & B0 --> e+ nu_e D*- gamma & 0
719         2 & anti-B0 --> mu- anti-nu_mu D*+ & 0
720         3 & D*- --> pi- anti-D0 & 1
721         4 & D*+ --> pi+ D0 & 2
722         5 & anti-D0 --> pi0 pi- K+ & 3
723         6 & D0 --> pi0 pi+ K- & 4
  
```

```

724
725 0 & Upsilon(4S) --> B0 anti-B0 & -1 & 2ndDcyTrInTb2
726 1 & B0 --> pi0 pi+ pi- rho- D- & 0
727 2 & anti-B0 --> mu- anti-nu_mu D*+ & 0
728 3 & rho- --> pi0 pi- & 1
729 4 & D- --> pi- pi- K+ & 1
730 5 & D*+ --> pi+ D0 & 2
731 6 & D0 --> K_L0 pi+ pi- & 5
732 }
733

```

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	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow e^+ \nu_e D^{*-} \gamma^F, \bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}, D^{*-} \rightarrow \pi^- \bar{D}^0,$ $D^{*+} \rightarrow \pi^+ D^0, \bar{D}^0 \rightarrow \pi^0 \pi^- K^+, D^0 \rightarrow \pi^0 \pi^+ K^-$ $(\Upsilon(4S) \dashrightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu \pi^0 \pi^0 \pi^+ \pi^+ \pi^- \pi^- K^+ K^- \gamma^F)$	0	3	3
2	$\Upsilon(4S) \rightarrow B^0 \bar{B}^0, B^0 \rightarrow \pi^0 \pi^+ \pi^+ \rho^- D^-, \bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}, \rho^- \rightarrow \pi^0 \pi^-,$ $D^- \rightarrow \pi^- \pi^- K^+, D^{*+} \rightarrow \pi^+ D^0, D^0 \rightarrow K_L^0 \pi^+ \pi^-$ $(\Upsilon(4S) \dashrightarrow \mu^- \bar{\nu}_\mu \pi^0 \pi^0 K_L^0 \pi^+ \pi^+ \pi^+ \pi^- \pi^- \pi^- K^+)$	1	2	5

4.2. Decay initial-final states

In a few cases, we have an 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. The obtained topology map is displayed in Table 13. The results are identical to those shown in the first two rows of Table 3.

```

736 % Signal identification — decay initial-final states
737 {
738     Y(4S) --> mu+ nu_mu 3 pi0 3 pi+ 4 pi- K+ K- & 2ndDcyIFStsInTb3
739     Y(4S) --> 5 pi0 5 pi+ 5 pi- K+ K- & 2ndDcyIFStsInTb3
740 }
741

```

Table 13: Signal decay initial-final states.

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

```

752 % Signal identification — particles
753

```

```

759   {
760     D*+  Dsp
761     J/psi Jpsi
762   }

```

763
764 Table 14 shows the resulting topology map. As a cross-check, the number of D^{*+} s in the table
765 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

766 4.4. Decay branches

767 On some occasions, we have to identify certain regular decay branches. Below is an ex-
768 ample demonstrating the related item by taking the two decay branches $\bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}$ and
769 $B^0 \rightarrow K_S^0 J/\psi$ as signals. Since regular decay branches also look like inclusive decay branches,
770 except that only two columns are involved in the item, the format of the input to the item for reg-
771 ular decay branches is identical to that for the component analysis over inclusive decay branches,
772 which is introduced in Section 3.8.

```

773
774 % Signal identification — decay branches
775 {
776   B0 --> mu- anti-nu_mu D*+ & B2munuDsp
777   B0 --> K_S0 J/psi & B2KsJpsi
778 }

```

779
780 The obtained topology map is displayed Table 15. For cross-checks, we note that the number of
781 $\bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}$ ($B^0 \rightarrow K_S^0 J/\psi$) in the table is equal to that in the first row of Table 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

782 4.5. Cascade decay branches

783 Sometimes, we are interested in certain cascade decay branches. The following example
784 shows the associated item with the two cascade decay branches $B^0 \rightarrow D^{*-} D_s^{*+}$, $D^{*-} \rightarrow \pi^- \bar{D}^0$,
785 $D_s^{*+} \rightarrow D_s^+ \gamma$ and $B^0 \rightarrow D^{*-} D_s^{*+}$, $D^{*-} \rightarrow \pi^- \bar{D}^0$ set as signals. While the first cascade decay
786 branch is identical to the fifth one in Table 7, the second is only part of it, which demonstrates
787 that the cascade decay branches supported in the item are not necessarily fully specified at the
788 level of certain hierarchy. Similar to decay trees, cascade decay branches are made up of regular
789 decay branches. Hence, the format of the input to the item for cascade decay branches is identical
790 to that for decay trees, which is introduced in Section 4.1.

```

791
792 % Signal identification — cascade decay branches
793 {
794   0 & B0 --> D*- D_s*+ & -1
795   1 & D*- --> pi- anti-D0 & 0

```

```

796      2 & D_s*+ --> D_s+ gamma & 0
797
798      0 & B0 --> D*- D_s*+ & -1
799      1 & D*- --> pi- anti-D0 & 0
800  }

```

801
802 Table 16 shows the resulting topology map. As a cross-check, the number of cases of the first
803 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	iSigCascDcyBrs	nCase	nCCase
1	$B^0 \rightarrow D^{*-} D_s^{*+}, D_s^{*-} \rightarrow \pi^- \bar{D}^0, D_s^{*+} \rightarrow D_s^+ \gamma$	0	1119	1119
2	$B^0 \rightarrow D^{*-} D_s^{*+}, D_s^{*-} \rightarrow \pi^- \bar{D}^0$	1	1180	2299

804 4.6. Inclusive decay branches

805 In a few cases, we have to identify some inclusive decay branches. Below is an example
806 demonstrating the related item by taking the two inclusive decay branches $\bar{B}^0 \rightarrow D^{*+} + \text{anything}$
807 and $B^0 \rightarrow K_S^0 + \text{anything}$ as signals. Except that only two columns are involved in the item, the
808 format of the input to the item is identical to that for the component analysis over inclusive decay
809 branches, which is introduced in Section 3.8.

```

810
811      % Signal identification — inclusive decay branches
812  {
813      anti-B0 --> D*+ & B2Dsp
814      B0 --> K_S0 & B2Ks
815  }

```

816
817 The obtained topology map is displayed in Table 17. As a cross-check, the number of $B^0 \rightarrow$
818 $K_S^0 + \text{anything}$ 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

819 4.7. Inclusive cascade decay branches

820 Occasionally, we may have an interest in certain inclusive cascade decay branches. The
821 following example shows the associated item with the two inclusive cascade decay branches
822 $\bar{B}^0 \rightarrow D^{*+} + \text{anything}$, $D^{*+} \rightarrow \pi^+ D^0$ and $B^0 \rightarrow K_S^0 J/\psi$, $K_S^0 \rightarrow \pi^+ \pi^-$, $J/\psi \rightarrow \mu^+ + \text{anything}$ set
823 as signals. Similar to decay trees and cascade decay branches, inclusive cascade decay branches
824 are made up of regular decay branches. Hence, the format of the input to the item for inclusive
825 cascade decay branches is also identical to that for decay trees, which is introduced in Section
826 4.1. and the independent textual name “*” denotes anything.

```

827
828      % Signal identification — inclusive cascade decay branches
829  {
830      0 & anti-B0 --> D*+ * & -1
831      1 & D*+ --> pi+ D0 & 0
832

```

```

833      0 & B0 --> K_S0 J/psi & -1
834      1 & K_S0 --> pi+ pi- & 0
835      2 & J/psi --> mu+ * & 0
836    }

```

837 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

839 4.8. Intermediate-resonance-allowed decay branches

840 On some occasions, we need to identify certain IRA decay branches. Below is an example
841 demonstrating the related item by taking the two IRA decay branches $D^{*+} \dashrightarrow \pi^0 \pi^+ \pi^+ K^-$ and
842 $J/\psi \dashrightarrow \pi^0 \pi^+ \pi^-$ as signals. Except that only two columns are involved in the item, the format
843 of the input to the item is identical to that for the component analysis over IRA decay branches,
844 which is introduced in Section 3.9.

```

845 % Signal identification — intermediate-resonance-allowed decay branches
846 {
847   D*+ --> K- pi+ pi+ pi0 & Dsp2K3Pi
848   J/psi --> pi+ pi- pi0 & Jpsi23Pi
849 }
850

```

851 The obtained topology map is displayed in Table 19. For the purpose of cross-checks, we
852 note that the number of $D^{*+} \dashrightarrow \pi^0 \pi^+ \pi^+ K^-$ 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	$D^{*+} \dashrightarrow \pi^0 \pi^+ \pi^+ K^-$	0	4971	4971
2	$J/\psi \dashrightarrow \pi^0 \pi^+ \pi^-$	1	59	5030

854 4.9. Essential topology tags

855 Table 20 summarizes and explains all of the essential topology tags involved in the signal
856 identification functionalities. For signal decay trees and signal decay initial-final states, there are
857 two sorts of topology tags. The first sort of tags, iSigDcyTr and iSigDcyIFSts, record the default
858 indices of the specified signal decay trees and signal decay initial-final states. They have similar
859 interpretations as iDcyTr and iDcyIFSts, and are shown in the third columns of Tables 12 and
860 13. The second sort of tags, nameSigDcyTr and nameSigDcyIFSts, save the specified aliases of
861 the signal decay trees and signal decay initial-final states. In cases the aliases are not specified,
862 empty strings will be stored.

863 For the latter six kinds of signal identification, there is only one sort of topology tags, which
864 records the number of instances of certain specified particle or decay branch found in each event.
865 Similar to the cases in the latter seven kinds of component analysis, in the topology tags, “i” in
866 “_i” is the default index of the specified particle or decay branch, and it ranges from 0 (included)

867 to the number of specified particles or decay branches (excluded). If the alias of the particle or
 868 decay 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 type	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; _i s
Decay branches	nSigDcyBr_i	number of signal decay branch; _i es
Cascade decay branches	nSigCascDcyBr_i	number of signal cascade decay branch; _i es
Inclusive decay branches	nSigIncDcyBr_i	number of signal inclusive decay branch; _i es
Inclusive cascade decay branches	nSigIncCascDcyBr_i	number of signal inclusive cascade decay branch; _i es
IRA decay branches	nSigRADcyBr_i	number of signal IRA decay branch; _i es

869 5. Common settings

870 From Sections 3 and 4, the optional parameters of the functionality items give us more choices
 871 and thus help us do our jobs quicker and better. In addition to these parameters, many optional
 872 items are designed and implemented to control the execution of the program in order to meet
 873 practical needs. Unlike the optional parameters, which only affect the individual functionalities
 874 to which they belong, the optional items have an impact on all of the functionalities, or at
 875 least most of the functionalities. The current version of the program contains 24 commonly used
 876 items, which can be divided into the following three groups: items on the input of the program,
 877 items on the functionalities of the program, and items on the output of the program. This section
 878 introduces these items in the three groups, with each group in one subsection.

879 5.1. Settings on the input of the program

880 5.1.1. Input entries

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

```
889 % Maximum number of entries to be processed
890 {
891   2000
892 }
```

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

```
896 % Cut to select entries
897 {
898 }
```

```

901      (X >-1) && (X <1)
902    }
903

```

904 Notably, only a single-line selection requirement is supported in the item, like the cases in the
905 methods Draw() [16] and GetEntries() [17] of the class TTree. In spite of this, such a requirement
906 is able to express any requirement with the help of the parentheses “()” as well as the logical
907 symbols “&&”, “||”, and “!”.

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

```

914
915 % Method to apply cut to array variables (Two options: T and F. Default: F)
916 {
917   T
918 }
919

```

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

921 5.1.2. Input decay branches

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

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

922 Normally, the program deals with all of the decay branches in every decay tree. However,
923 examining all the branches is not always required in practice. Sometimes, we only concern the
924 first n hierarchies of the branches. Similar to that in cascade decay branches of particles (as we
925 introduce in Section 3.6), the hierarchy here reflects the rank of a decay branch in a decay tree.
926 For example, in the decay tree $\Upsilon(4S) \rightarrow B^0 \bar{B}^0$, $B^0 \rightarrow e^+ \nu_e D^{*-} \gamma^F$, $\bar{B}^0 \rightarrow \mu^- \bar{\nu}_\mu D^{*+}$, $D^{*-} \rightarrow \pi^- \bar{D}^0$,
927 $D^{*+} \rightarrow \pi^+ D^0$, $\bar{D}^0 \rightarrow \pi^0 \pi^- K^+$, $D^0 \rightarrow \pi^0 \pi^+ K^-$, the hierarchies of the seven individual branches
928 are 1, 2, 2, 3, 3, 4, and 4, respectively. The program provides an item to set the maximum hierarchy.
929 Below is an example showing the item with the maximum hierarchy set at one.

```

930
931 % Maximum hierarchy of heading decay branches to be processed in each event
932 {
933   1
934 }
935

```

936 With the setting, the decay branches with hierarchy larger than one will be ignored by the
937 program. For the component analysis over the decay trees of the $\Upsilon(4S)$ sample, only the first hierar-
938 chy of $\Upsilon(4S)$ decay branches are analyzed, and the result is shown in Table 21. From the table,
939 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

940 mixing.
941 Similarly, in the case of the maximum hierarchy set at two, we could get the result of the com-
942 ponent analysis over the first two hierarchies of $\Upsilon(4S)$ decay branches, as displayed in Table
943 22.

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

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

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

```
947
948 % Ignore the decay of the following particles
949 {
950   B0
951   anti-B0
952 }
```

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

```
957
958 % Ignore the decay of the daughters of the following particles
959 {
960   B0
961   anti-B0
962 }
```

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

966 As mentioned in Section 2.6, the decay $\pi^0 \rightarrow \gamma\gamma$ is ignored by default. On the occasions
967 when we need to identify the signals involving the decay, we can make the program retain the
968 decay with the item below set to “Y”.

```
969
970 % Retain the decay of pi0 to gamma gamma (Two options: Y and N. Default: N)
971 {
972   Y
973 }
```

976
 977 Besides, if needed, one can make the program ignore other final decay branches, such as $\eta \rightarrow \gamma\gamma$
 978 and $K_S^0 \rightarrow \pi^+\pi^-$, with the following item.
 979

```

980   % Ignore the following final decay branches
981   {
982     eta --> gamma gamma
983     K_S0 --> pi+ pi-
984   }
  
```

985 5.1.3. Initial and final state radiation photons

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

```

993   % PDG code of ISR photons (Default: 222222222)
994   {
995     222222222
996   }
997
998
999   % PDG code of FSR photons (Default: -22)
1000  {
1001    -22
1002  }
1003
  
```

1004 In this case, the program is able to label the ISR and FSR photons as γ^i (gammaI) and γ^f (gam-
 1005 maf) in the output pdf (plain text) files, respectively.

1006 On other occasions, ISR and FSR photons are not marked out in advance due to some reasons.
 1007 In such cases, the program has to identify them by itself according to the following rules: photons
 1008 who have no mothers recorded in the arrays of the PDG codes and mother indices are considered
 1009 as generalized ISR photons, while other photons who have at least one $e^\pm, \mu^\pm, \pi^\pm, K^\pm, p$, or \bar{p}
 1010 sister are taken as generalized FSR photons. Here, the modifier “generalized” is used because
 1011 the rules can not determine the types of the photons in absolute accuracy. For example, photons
 1012 from radiative decays might be mistaken as FSR photons. Despite this, generalized ISR and FSR
 1013 photons are good concepts, particularly in cases where the sources of the photons are not required
 1014 to be distinguished clearly. The program will label the generalized ISR and FSR photons as γ^I
 1015 (gammaI) and γ^F (gammaF) in the output pdf (plain text) files, respectively.

1016 Notably, we are not concerned about these ISR and FSR photons in many cases, particularly
 1017 when we want to identify our signals from some samples. If they have already been marked out
 1018 beforehand, one can make the program ignore them accurately by setting the following two items
 1019 to “Ys”.

```

1021   % Ignore ISR photons (Three options: Ys, Yg and N. Default: N)
1022   {
1023     Ys
1024   }
1025
1026
1027   % Ignore FSR photons (Three options: Ys, Yg and N. Default: N)
1028
  
```

```

1029    {
1030        Ys
1031    }
1032
1033 In cases that these photons are not marked in advance, the option "Yg" can be used to ignore the
1034 generalized ISR and FSR photons. In "Ys" and "Yg", "s" and "g" are the initials of the words
1035 "strict" and "generalized", respectively.

```

1036 *5.2. Settings on the functionalities of the program*

1037 *5.2.1. Candidate based analysis*

1038 According to the number of signal candidates in an event that are selected and retained to
1039 extract physics results, data analysis in high energy experiments can be divided into the following
1040 two categories: event based analysis and candidate based analysis. While at most one candidate
1041 in an event is kept in event based analysis, one or more candidates in an event can be retained in
1042 candidate based analysis. Generally, the quantities related to a candidate are stored in an entry of
1043 the TTree objects in the root files. Thus, one or more entries relate to an event in candidate based
1044 analysis, while only one entry corresponds to an event in event based analysis. Normally, the
1045 indices of candidates within an event are stored in the corresponding entries in candidate based
1046 analysis.

1047 By default, the program analyzes the input entries one by one. In this case, the events with
1048 multiple candidates will be processed repeatedly. Particularly, the number of physics processes at
1049 the sample level will be overcounted. One can make the program avoid the problem by inputting
1050 "Y" to the following item.

```

1051     % Avoid over counting for candidate based analysis (Two options: Y and N. Default: N)
1052     {
1053         Y
1054     }
1055

```

1056 Also, the indices of candidates within an event are required. We can tell the program the related
1057 TBranch name with the following item.

```

1058
1059     % TBranch name of the indices of candidates in an event (Default: __candidate__)
1060     {
1061         iCandidate
1062     }
1063

```

1064 With the settings, the program will process the first entry of each event in a normal way, including
1065 obtaining and storing the topology tags; it will not analyze the other entries of the same event,
1066 but only store the same topology tags to them.

1068 *5.2.2. Charge conjugation*

1069 Charge conjugation is an important concept in high energy physics. By default, charge con-
1070 jugate objects (particles and decays) are processed separately in the program. However, we need
1071 to handle them together in many physics studies because of the sameness between them. One
1072 can have the program process them together with the item below set to "Y".

```

1073
1074     % Process charge conjugate objects together (Two options: Y and N. Default: N)
1075     {
1076         Y
1077

```

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. For self-charge-conjugate objects (particles or decays), the charge conjugate indices have the value 0; for non-self-charge-conjugate objects, they have the value 1 or -1 ; while 1 tags the objects presented in the topology maps, -1 indicates their charge conjugate objects.

Component type	Topology tag	Interpretation
Decay trees	iCcDcyTr	index_{cc} of decay tree
Decay initial-final states	iCcDcyIFsts	index_{cc} of decay initial-final states
	iCpDcyBr.i	index_{cc} of particle _i
Decay branches of particles	(1) iCcDcyBrP.i.j (2) nCcPDcyBr.i (2) iDcyBrCcP.i.j (2) nAllPDcyBr.i	index_{cc} of decay branch of the j th particle _i number of cc particle _i s (decay branches) index of decay branch of the j th cc particle _i number of all particle _i s (decay branches)
Production branches of particles	iCpProdBr.i (1) iCcProdBrP.i.j (2) nCcPProdBr.i (2) iProdBrCcP.i.j (2) nAllPProdBr.i	index_{cc} of particle _i index_{cc} of production branch of the j th particle _i number of cc particle _i s (production branches) index of production branch of the j th cc particle _i number of all particle _i s (production branches)
Mothers of particles	iCpMoth.i (1) iCcMothP.i.j (2) nCcPMoth.i (2) PDGMothCcP.i.j (2) nAllPMoth.i	index_{cc} of particle _i index_{cc} of mother of the j th particle _i number of cc particle _i s (mothers) PDG code of mother of the j th cc particle _i number of all particle _i s (mothers)
Cascade decay branches of particles	iCpCascDcyBr.i (1) iCcCascDcyBrP.i.j (2) nCcPCascDcyBr.i (2) iCascDcyBrCcP.i.j (2) nAllPCascDcyBr.i	index_{cc} of particle _i index_{cc} of cascade decay branch of the j th particle _i number of cc particle _i s (cascade decay branches) index of cascade decay branch of the j th cc particle _i number of all particle _i s (cascade decay branches)
Decay final states of particles	iCpDcyFSt.i (1) iCcDcyFStP.i.j (2) nCcPDcyFSt.i (2) iDcyFStCcP.i.j (2) nAllPDcyFSt.i	index_{cc} of particle _i index_{cc} of decay final state of the j th particle _i number of cc particle _i s (decay final states) index of decay final state of the j th cc particle _i number of all particle _i s (decay final states)
Inclusive decay branches	iCpIncDcyBr.i (1) iCcDcyBrIncDcyBr.i.j (2) nCcIncDcyBr.i (2) iDcyBrCcIncDcyBr.i.j (2) nAllIncDcyBr.i	index_{cc} of inclusive decay branch _i index_{cc} of decay branch of the j th inclusive decay branch _i number of cc inclusive decay branch _i s index of decay branch of the j th cc inclusive decay branch _i number of all inclusive decay branch _i s
IRA decay branches	iCpRADcyBr.i (1) iCcDcyBrIRADcyBr.i.j (2) nCcIRADcyBr.i (2) iDcyBrCcIRADcyBr.i.j (2) nAllIRADcyBr.i	index_{cc} of IRA decay branch _i index_{cc} of decay branch of the j th IRA decay branch _i number of cc IRA decay branch _i s index of decay branch of the j th cc IRA decay branch _i number of all IRA decay branch _i s

1078 }
 1079
 1080 Performing topology analysis with this setting inserts new topology tags in the output root files
 1081 and adds new counters to topology maps in the output plain text, tex source, and pdf files. Tables
 1082 23 and 24 list and interpret all of the topology tags related to charge conjugation involved in the
 1083 component analysis and signal identification functionalities, respectively.

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. For self-charge-conjugate objects (particles or decays), the charge conjugate indices have the value 0; for non-self-charge-conjugate objects, they have the value 1 or -1 : while 1 tags the objects presented in the topology maps, -1 indicates their charge conjugate objects.

Signal type	Topology tag	Interpretation
Decay trees	$i\text{CcSigDcyTr}$	index_{cc} of signal decay tree
Decay initial-final states	$i\text{CcSigDcyIFSts}$	index_{cc} of signal decay initial-final states
Particles	$i\text{CcSigP.i}$ (*) $n\text{CcSigP.i}$ (*) $n\text{AllSigP.i}$	index_{cc} of signal particle; number of cc signal particle; number of all signal particle;
Decay branches	$i\text{CcSigDcyBr.i}$ (*) $n\text{CcSigDcyBr.i}$ (*) $n\text{AllSigDcyBr.i}$	index_{cc} of signal decay branch; number of cc signal decay branch; number of all signal decay branch;
Cascade decay branches	$i\text{CcSigCascDcyBr.i}$ (*) $n\text{CcSigCascDcyBr.i}$ (*) $n\text{AllSigCascDcyBr.i}$	index_{cc} of signal cascade decay branch; number of cc signal cascade decay branch; number of all signal cascade decay branch;
Inclusive decay branches	$i\text{CcSigIncDcyBr.i}$ (*) $n\text{CcSigIncDcyBr.i}$ (*) $n\text{AllSigIncDcyBr.i}$	index_{cc} of signal inclusive decay branch; number of cc signal inclusive decay branch; number of all signal inclusive decay branch;
Inclusive cascade decay branches	$i\text{CcSigIncCascDcyBr.i}$ (*) $n\text{CcSigIncCascDcyBr.i}$ (*) $n\text{AllSigIncCascDcyBr.i}$	index_{cc} of signal inclusive cascade decay branch; number of cc signal inclusive cascade decay branch; number of all signal inclusive cascade decay branch;
IRA decay branches	$i\text{CcSigIRADcyBr.i}$ (*) $n\text{CcSigIRADcyBr.i}$ (*) $n\text{AllSigIRADcyBr.i}$	index_{cc} of signal IRA decay branch; number of cc signal IRA decay branch; number of all signal IRA decay branch;

1084 As an example, we perform the component analysis over decay trees with the charge con-
 1085 jugate item. Table 25 shows the obtained topology map. Besides the columns in Table 2, two
 1086 additional columns with the headers “nCcEtr” and “nAllEtr” are inserted in the table. Here, “nC-
 1087 cEtr” represents the number of entries involving the charge conjugate decay trees, and “nAllEtr”
 1088 is the sum of “nEtr” and “nCcEtr”. In addition to “iDcyTr”, “iCcDcyTr” is also inserted in the
 1089 output root files as a topology tag. It is short for charge conjugate index of decay tree. For self-
 1090 charge-conjugate decay trees, it has the value 0; for non-self-charge-conjugate decay trees, it has
 1091 the value 1 or -1 : while 1 tags the decay trees listed in the topology maps, -1 indicates their
 1092 charge conjugate decay trees. Whereas the equal values of “iDcyTr” for each decay tree and its
 1093 charge conjugate decay tree indicate their sameness, the opposite values of “iCcDcyTr” for them
 1094 reflect their difference.

Table 25: Decay trees and their respective initial-final states (with the charge conjugation setting).

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

1095 As another example, we carry out the component analysis over the decay branches of D^{*+}
 1096 and J/ψ . The resulting topology map of D^{*+} is displayed in Table 26. Compared with Table
 1097 4, two new columns are added to the table, and their headers “nCcCase” and “nAllCase” have
 1098 similar meanings as “nCcEtr” and “nAllEtr” in Table 25. For a specified particle, what we want
 1099 to further record with topology tags are as follows: (1) whether it is self-charge-conjugate;
 1100 (2) whether its decay branches are self-charge-conjugate, if it is self-charge-conjugate; (3) the
 1101 number and the indices of the decay branches of its charge-conjugate particle, if it is not self-
 1102 charge-conjugate. Hence, in addition to “nPDcyBr.i” and “iDcyBrP.i.j”, the following topology
 1103 tags are also inserted in the output root files: “iCcPDcyBr.i” for all specified particles; “iCcD-
 1104 cyBrP.i.j” for self-charge-conjugate particles only; and “nCcPDcyBr.i”, “iDcyBrCcP.i.j”, and
 1105 “nAllPDcyBr.i” for non-self-charge-conjugate particles only. Here, “iCcPDcyBr.i” tags whether
 1106 the i^{th} particle is self-charge-conjugate. For self-charge-conjugate particles, it has the value 0;
 1107 for non-self-charge-conjugate particles, it has the value 1.

 Table 26: Decay branches of D^{*+} (with the charge conjugation setting).

rowNo	decay branch of D^{*+}	iDcyBrP	nCase	nCcCase	nAllCase	nCCase
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

1108 The topology tag “iCcDcyBrP.i.j” records the charge conjugation property of the decay
 1109 branch of the j^{th} instance of the i^{th} particle. It is to “iDcyBrP.i.j” what “iCcDcyTr” is to “iD-
 1110 cyTr”. The topology tag “iDcyBrCcP.i.j” is designed for the charge conjugate particle of the i^{th}
 1111 particle (for D^{*-} in this example). It has a similar meaning as “iDcyBrP.i.j”. Particularly, the
 1112 values of “iDcyBrP.i.j” and “iDcyBrCcP.i.j” tagging charge conjugate decay branches are equal
 1113 to each other. The topology tag “nCcPDcyBr.i” stands for the number of the charge conjugate

```
1114 ith particles (or their decay branches) found in each event, and “nAllPDcyBr_i” is the sum of  
1115 “nPcyBr_i” and “nCcPDcyBr_i”.
```

1116 5.2.3. *Settings only on signal identification*

1117 Normally, the signals specified in the signal identification functionality items are both tagged
1118 and counted by executing the program one time. In the case of a huge sample that will take a long
1119 time, it is a good idea to first tag the signals with multiple jobs each running on one machine, and
1120 then count the tagged signals together. One can make the program carry out the idea by setting
1121 the following item to “T” and “C” in the first and second steps, respectively. Here, “T” and “C”
1122 stand for tagging and counting, respectively.

```
1123 % Analysis tasks for signal identifications (Three options: TC, T and C. Default: TC)  
1124 {  
1125     T  
1126 }  
1127
```

1128 By default, the signals set in the signal identification functionality items are listed in the output
1129 plain text, tex source, and pdf files in the sequence they are specified. In cases of plenty of
1130 signals, there is probably a need to sort them according to the number of cases found in the input
1131 samples. One can have the program do the sorting by inputting “Y” to the item below.

```
1132 % Sort the signals in the topology maps related to signal identifications (Two options: Y and N. Default: N)  
1133 {  
1134     Y  
1135 }  
1136
```

1138 5.3. *Settings on the output of the program*

1139 By default, decay objects (trees, initial-final states, and branches) are left-aligned in the output
1140 pdf files. If one likes it, he/she can request the program to center them by setting the following
1141 item to “Y”.

```
1142 % Center decay objects in output pdf files (Two options: Y and N. Default: N)  
1143 {  
1144     Y  
1145 }  
1146
```

1147 As mentioned in Section 2.6, after the execution of the program, one or more root files will
1148 be output to save topology tags. By default, the program switches to a new output file whenever
1149 the size of the TTree object in memory exceeds 3 GB. In addition to this, the program provides
1150 an item to control the switch of output files by setting the maximum number of entries to be
1151 saved in a single output file. The following example shows the item with the maximum number
1152 set to 1 million.

```
1153 % Maximum number of entries to be saved in a single output root file  
1154 {  
1155     1000000  
1156 }  
1157
```

1158 Besides, one can have the program generate one output file by one input file with the following
1159 item set to “Y”.

```
1160 % One output root file by one input root file (Two options: Y and N. Default: N)  
1161 {  
1162     Y  
1163 }  
1164
```

1165 In default cases, flat TBranch objects are used to store topology tags in the output root files.

1169 This is necessary for the Belle II experiment, as array TBranch objects are not recommended to
 1170 use in physics analyses in order to use other tools such as NumPy [11] and pandas [12]. However,
 1171 since array TBranch objects are elegant and efficient in organizing and storing homogeneous
 1172 data, sometimes it is better to use them than flat TBranch objects in other experiments, such as
 1173 the BESIII experiment. One can make the program use array TBranch objects to store topology
 1174 tags by inputting “Y” to the item below.

```

1175     % Use array tbranches to store topology tags in output root files when possible (Two options: Y and N. Default: N)
1176     {
1177         Y
1178     }
1180
  
```

1181 By default, to facilitate the validation of topology analysis results, the input TBranch objects
 1182 are copied to the output root files along with other TBranch objects for physics analyses. How-
 1183 ever, they often occupy too much disk space and are useless for following physics analyses. In
 1184 the case of being flat, a massive amount of these TBranch objects also looks awkward. Thus,
 1185 after the validation with a small sample, it would be better to remove these TBranch objects from
 1186 the output root files. One can request the program to perform this removal operation before it
 1187 terminates by setting the following item to “Y”.

```

1188     % Remove the input tbranches from output root files (Two options: Y and N. Default: N)
1189     {
1190         Y
1191     }
1193
  
```

1194 In all of the previous examples, the program is applied to the inclusive MC samples in e^+e^-
 1195 colliding experiments. Besides, the program can also be used in other types of high energy ex-
 1196 periments, for example, the PANDA experiment [18], a $p\bar{p}$ annihilation experiment under con-
 1197 struction at Darmstadt, Germany. On these occasions, we have to specify the right initial state
 1198 particles with the following item to obtain the proper topology maps.

```

1199     % Initial state particles (Default: e- e+)
1200     {
1201         anti-p- p+
1203     }
1204
  
```

1205 With the setting, the default initial state e^+e^- is replaced by $p\bar{p}$, as shown in Table 27, which
 1206 displays the results of a component analysis over decay trees of a small $p\bar{p}$ annihilation sample.

Table 27: Decay trees and their respective initial-final states ($p\bar{p}$ annihilation).

rowNo	decay tree	decay final state	iDcyTr	nEtr	nCEtr
1	$p\bar{p} \rightarrow p\bar{p}$	$p\bar{p}$	1	232	232
2	$p\bar{p} \rightarrow \pi^+\pi^- p\bar{p}$	$\pi^+\pi^- p\bar{p}$	24	53	285
3	$p\bar{p} \rightarrow \pi^0 p\bar{p}$	$\pi^0 p\bar{p}$	5	35	320
4	$p\bar{p} \rightarrow \pi^0\pi^+\pi^- p\bar{p}$	$\pi^0\pi^+\pi^- p\bar{p}$	0	33	353
5	$p\bar{p} \rightarrow \pi^0\pi^0\pi^0\pi^+\pi^-\pi^-$	$\pi^0\pi^0\pi^+\pi^-\pi^-$	39	31	384
rest	$p\bar{p} \rightarrow$ others (184 in total)	corresponding to others	—	616	1000

1207 6. Auxiliary facilities

1208 This section introduces some auxiliary facilities for the use of the program, including a card
 1209 file to preset frequently used items; some additional command line arguments to reset the names

1210 of input root files, the common name of output files, and the maximum number of entries to be
1211 processed; and two commands implemented in tex source files. Different from that presented in
1212 the previous four sections, the content presented in this section is not the essential part of the
1213 program. However, with these auxiliary facilities, we can make the program do our jobs better
1214 and quicker on some occasions.

1215 *6.1. The underlying card file*

1216 A card file, namely “underlying_topoana.card” under the directory “share”, to preset frequently
1217 used items is developed to assist the card file specified by the first argument of the
1218 command “topoana.exe”. Here, we refer to the former and latter card files as underlying and
1219 primary, respectively. In general, the primary card file is sufficient to set items for the execution
1220 of the program. However, considering some items are frequently used with constant inputs by a
1221 user or a group of users, it is better to move the items from the primary card file to the underlying
1222 card file, in order to make the primary card file more concise and make us more focused on the
1223 items specially set for the dedicated topology analysis.

1224 One can decide whether to set an item in the underlying card file according to his/her own
1225 needs. Here, we introduce some frequently used items that are suitable to be put in the underlying
1226 card file as follows. As mentioned in Section 2.4, the items related to the storage type and
1227 TBranches names of the input data are usually fixed for a user or a group of users. Thus, it is
1228 quite appropriate to move them to the underlying card file. We have to process charge conjugation
1229 particles and decays together in many physics studies. In such studies, it is also a good practice
1230 to put the item on charge conjugation in the underlying card file.

1231 The program first reads the items in the underlying card file and then reads those in the
1232 primary card file. The items set in the underlying card file can be reset in the primary card file.
1233 In such a case, the inputs in the underlying card file will be replaced by their counterparts in the
1234 primary card file.

1235 *6.2. Additional command line arguments*

1236 Normally, only the “cardFileName” is required to be passed as an argument of the command
1237 “topoana.exe”, and all of the necessary information can be configured via the setting items filled
1238 in the card file. On some occasions, we need to run the program over multiple samples separately,
1239 with identical settings except for the names of input root files and the common name of output
1240 files. A regular approach to do such a job requires multiple card files, each corresponding to
1241 one sample. This approach appears a bit tedious in cases of many samples. To avoid this, two
1242 additional command line arguments are designed and implemented to reset the names of input
1243 root files and the common name of output files. Similarly, an argument is also developed for the
1244 maximum number of entries to be processed.

1245 These optional arguments should be typed with prompts, which are listed and explained as
1246 follows.

- 1247 • **-i:** The names of input root files should be provided after the prompt. One or more names
1248 are allowed here. They will replace those set in the card file.
- 1249 • **-t:** The TTree name should be provided after the prompt. It will replace the one set in the
1250 card file.
- 1251 • **-o:** The common name of output files should be provided after the prompt. It will replace
1252 the one set in the card file or the default one, that is, the name of the card file.

- 1253 • –n: The maximum number of entries to be processed should be provided after the prompt.
 1254 It will replace that set in the card file.

1255 *6.3. Commands implemented in tex source files*

1256 The output pdf files can be checked after the execution of the program. If their styles are not
 1257 to our taste, we can edit the corresponding tex source files to get the desired styles, according
 1258 to the regular LaTeX rules. Besides the rules, two commands are implemented in the tex source
 1259 files to help us edit the files quickly and easily for two common desired styles.

1260 By default, topology tags are listed along with topology maps in the output plain text, tex
 1261 source, and pdf files. However, only the topology maps are needed on some occasions, especially
 1262 in presentations. In such cases, one can suppress the topology tags in the output tex source
 1263 and pdf files by simply changing the definition of the cmtTopoTags command from the nominal
 1264 one

1265
 1266 \newcommand{\topoTags}[1]{#1}

1267
 1268 to the alternative one

1269
 1270 \newcommand{\topoTags}[1]{}
 1271

1272 in the preamble of the text source files. Here, “#1” is the formal parameter of the string for
 1273 the topology tags. With the nominal definition, “\topoTags{#1}” returns the string exactly, while
 1274 with the alternative definition it only returns an empty string. That is why the definition below is
 1275 able to suppress the topology tags.

1276 After the revision of the tex source files, one can re-compile them with the pdflatex command.
 1277 Usually, the pdflatex command has to be executed two or three times for a fully compiled pdf
 1278 file, and many undesired files in other formats are generated during the compilation. To execute
 1279 the pdflatex command and remove the undesired files at one stroke, we develop a bash script,
 1280 namely “getPdfFromTex.sh” under the directory “utilities”. The script should be executed with
 1281 the following command line: `getPdfFlFromTexFl.sh texFileName`. Compiling the tex source
 1282 files with the script is recommended.

1283 *7. Summary*

1284 We develop a program, namely TopoAna, with C++, ROOT, and LaTeX for the event type
 1285 analysis of inclusive MC samples in high energy physics experiments. This user guide provides
 1286 a detailed description of the program, including a basic introduction to it, two categories of its
 1287 functionalities — component analysis and signal identification, and some common settings and
 1288 auxiliary facilities for its execution. The program has rich functionalities and aims to solve all
 1289 kinds of event type analysis tasks. Meanwhile, it is easy to use and has a high processing rate.
 1290 These features make the program a powerful tool to analyze the backgrounds involved in our
 1291 research works and to identify the physics processes of interests from the inclusive MC samples.

1292 Since it does not rely on any specific software frameworks, the program applies to many high
 1293 energy physics experiments. Up to now, it has been put into use in three experiments at e^+e^-
 1294 colliders: the BESIII, Belle, and Belle II experiments. Besides these experiments, it can also be
 1295 used in other types of experiments, such as the PANDA experiment, a $p\bar{p}$ annihilation exper-
 1296 iment. Also, the program is applicable to the future e^+e^- colliding experiments under research

1297 and development, such as the circular electron-positron collider (CEPC) [19, 20] experiment in
1298 China, the super Charm- τ factory (SCTF) experiment [21] in Russia, and the super τ -Charm fac-
1299 tory (STCF) experiment [22] in China. These experiments offer wide space for the application
1300 of the program.

1301 On the other hand, we note that the application of the program to some other experiments is
1302 limited. For example, thousands of particles can be produced from dozens of pp collisions in
1303 an event of the ATLAS [23] and CMS [24] experiments at the LHC [25]; in such cases, there
1304 is little point in performing the event type analysis of corresponding MC samples. Nonetheless,
1305 the application scope of the program is still broad. In particular, it applies to the e^+e^- colliding
1306 experiments where at most tens of particles are produced from the annihilation of a pair of e^+e^-
1307 in an event. With more user needs coming out in the future, we will further extend and perfect it
1308 to make it more powerful and well-rounded.

1309 Acknowledgements

1310 This work was supported by the National Natural Science Foundation of China [grant num-
1311 bers 11575017, 11661141008, 11761141009, 11875262, 11975076] and the CAS Center for
1312 Excellence in Particle Physics (CCEPP). In addition, we would like to thank all of the people
1313 who have helped us in the development of the program. We first thank Prof. Changzheng Yuan,
1314 Bo Xin, and Haixuan Chen for their help at the early stage of developing the program. We are
1315 particularly grateful to Prof. Xingtao Huang for his comments on the principles and styles of the
1316 program, to Remco de Boer for his suggestions on the tex output and the use of GitHub, and to Xi
1317 Chen for his discussions on the core algorithms. We are especially indebted to Prof. Xiqing Hao,
1318 Longke Li, Xiaoping Qin, Ilya Komarov, Yubo Li, Guanda Gong, Suxian Li, Junhao Yin, Prof.
1319 Xiaolong Wang, Yeqi Chen, and Hannah Wakeling for their advice in extending and perfecting
1320 the program. Also, we thank Xi'an Xiong, Runqiu Ma, Wencheng Yan, Sen Jia, Lu Cao, Dong
1321 Liu, Hongpeng Wang, Jiawei Zhang, Hongrong Qi, Jiajun Liu, Maoqiang Jing, Yi Zhang, Wei
1322 Shan, and Yadi Wang for their efforts in helping us test the program.

1323 References

- 1324 [1] ROOT User's Guide, Available online: <https://root.cern/root/html/doc/guides/users-guide/ROOTUsersGuide.html>.
- 1325 [2] Documentation of the TFile class, Available online: <https://root.cern/root/html534/TFile.html>.
- 1326 [3] K.T. Chao, Y.F. Wang, et al., *Int. J. Mod. Phys. A* 24 (2009) S1-794.
- 1327 [4] M. Ablikim, et al. (BESIII Collaboration), White Paper on the Future Physics Programme of BESIII, arXiv:1912.05983.
- 1328 [5] E. Kou, et al., *Prog. Theor. Exp. Phys.* 2019 (2019) 123C01.
- 1329 [6] J. Brodzicka, T. Browder, P. Chang, et al., *Prog. Theor. Exp. Phys.* 2012 (2012) 04D001.
- 1330 [7] Text of MIT license, Available online: <https://mit-license.org/>.
- 1331 [8] Documentation of the TTree class, Available online: <https://root.cern/root/html534/TTree.html>.
- 1332 [9] M. Tanabashi, et al. (Particle Data Group), *Phys. Rev. D* 98 (2018) 030001.
- 1333 [10] Documentation of the TBranch class, Available online: <https://root.cern/root/html534/TBranch.html>.
- 1334 [11] Documentation of NumPy, Available online: <https://numpy.org/devdocs/>.
- 1335 [12] Documentation of pandas, Available online: <https://pandas.pydata.org/pandas-docs/stable/>.
- 1336 [13] Documentation of the TChain class, Available online: <https://root.cern/root/html534/TChain.html>.
- 1337 [14] Reference of unordered maps, Available online: http://www.cplusplus.com/reference/unordered_map/unordered_map/.
- 1338 [15] D. J. Lange, *Nucl. Instrum. Meth. A* 462 (2001) 152.
- 1339 [16] Documentation of the Draw() method of the TTree class, Available online: <https://root.cern/root/html534/TTree.html#TTree:Draw@2>.
- 1340 [17] Documentation of the GetEntries() method of the TTree class, Available online: <https://root.cern/root/html534/TTree.html#TTree:GetEntries@1>.

- 1344 [18] W. Erni, et al. (PANDA Collaboration), Physics Performance Report for PANDA: Strong Interaction Studies with
1345 Antiprotons, arXiv:0903.3905.
1346 [19] CEPC CDR Volume 1 (Accelerator), Available online: http://cepc.ihep.ac.cn/CEPC_CDR_Vol1_Accelerator.pdf.
1347 [20] CEPC CDR Volume 2 (Physics & Detector), Available online: http://cepc.ihep.ac.cn/CEPC_CDR_Vol2_Physics-Detector.pdf.
1348 [21] A.E. Bondar, et al. (Charm-Tau Factory Collaboration), Phys. Atom. Nucl. 76 (2013) 1072.
1349 [22] Q. Luo, D. Xu, “Progress on Preliminary Conceptual Study of HIEPA, a Super Tau-Charm Factory in China”, in
1350 Proc. 9th International Particle Accelerator Conf. (IPAC2018), Vancouver, BC, Canada, 422.
1351 [23] G. Aad, et al. (ATLAS Collaboration), JINST 3 (2008) S08003.
1352 [24] S. Chatrchyan, et al. (CMS Collaboration), JINST 3 (2008) S08004.
1353 [25] L. Evans (ed.), P. Bryant (ed.), JINST 3 (2008) S08001.
1354