E-166 note, September 9, 2005

Analysis of positron asymmetries for June data

Andreas Schälicke, Hermann Kolanoski, Sabine Riemann

DESY Zeuthen

Abstract

In view of the just started new run period, a summary of the analysis of positron asymmetry for June 2005 data shall be given. In discussion with the analysis groups in Hamburg and Zeuthen, a new robust method to analyse the CsI data has been agreed on. With this method and using 104 data runs, corresponding approx. 6 hour run time, the asymmetry is found to be 0.0078 ± 0.0010 (statistical error only).

1 Introduction

In the data from the June run of the E166 experiment, 104 runs can be used for the determination of the positron asymmetry. To discriminate signal from background the undulator was turned off for every second electron bunch. Thus, each run contains approximately equal numbers of pure background events (B) and events with signal plus background ($\tilde{S}$). The background has to be subtracted from the signal events to get the pure signal.

For the analysis the central CsI crystal was used only, since all other crystals have much worse signal-background ratios[1].

In section 2, a new, robust method to combine signal and background measurements is introduced. Sec. 3 gives all necessary ingredients for the asymmetry determination. In sec. 4 the results of this analysis is summarised. An appendix, illustrating the properties of the new method with the help of a *Toy Monte Carlo* completes this note.

2 Obtaining a background subtraction for each run

2.1 How to make the best out of current data

When analysing the data for the positron signal, we observe a wide spread of calculated asymmetries, even when taking the same data and using similar methods. Although the observed differences are consistent within 2σ the given error estimates, this situation is not at all satisfying. One might expect the "true" value only to be within the range of the error, if different analyses of the same data yield different results one has to become alert.

For instance, the results appear to be very sensible to small changes in the following analysis setup:

- first data point used (staring with event 1 or event 2),
- method of pairing background and signal events,
- fit region, and binning of histograms,
- fit method χ^2 vs. likelihood.

This shall be quantified in the following.

[1] Taking the whole CsI detector actually decreases the significance of the result.

First, a simple method is considered to separated the signal by subtracting event by event the observed background.For instance Run 1227 contains 964 valid signal and 981 valid background events. These events can be used to pair adjacent $\tilde{S}$ and B events, but such a pairing is not unique. Indeed one can create two completely different pairings, just by performing the same analysis but starting with the second event[2]. Then it turns out that the obtained result may depend on the pairing (cf. plots at the end), for instance one pairing may return a mean $(\tilde{S} - B)$ of 0.706 ± 0.011, while the other gives 0.720 ± 0.011. In the end such small differences can sum up to quite different asymmetries. This dependence has to be understood and included in the error of the asymmetry. An analysis which does not depend on the "method" to select the signal-background pairs seems to be advantageous, and shall be illustrated in the following.

2.2 Proposed strategy

The aim of this new strategy is to avoid to use only one specific choice of pairing of signal and background events by taking a proper average over the background events (both, events before and after the signal event).

The basic idea shall be illustrated with an example to calculate the mean difference of two given distributions of signal and background events. The most simple way is to just calculate the difference of the mean of individual distributions, but there are other possibilities to obtain *exactly* the same result (as long as no cuts are applied).

Consider now different possibilities:

Method 1: Calculate mean of background and signal (+background) separately:

$$\begin{aligned} < S > &= \frac{1}{n} \sum_{i=1}^{n} s_i \ , \\ < B > &= \frac{1}{n} \sum_{i=1}^{n} b_i \ , \end{aligned} \tag{1}$$

and determine the difference:

$$< S - B > = < S > - < B > \ . \tag{2}$$

Method 2: Provide pairs of signal and background and evaluate its mean:

$$< S - B > = \frac{1}{n} \sum_{i=1}^{n} (s_i - b_i) \ . \tag{3}$$

Method 3: Similar as method 2 but calculate the mean background first

$$< S - B > = \frac{1}{n} \sum_{i=1}^{n} (s_i - < B >) \ . \tag{4}$$

Method 4: Use all possible pairs and calculate

$$< S - B > = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} (s_i - b_j) \tag{5}$$

or if the number of signal measurements n_s and the number of background measurements n_b do not coincide

$$< S - B > = \frac{1}{n_s} \frac{1}{n_b} \sum_{i=1}^{n_s} \sum_{j=1}^{n_b} (s_i - b_j) \tag{6}$$

[2]The maximal number of valid (non-overlapping) pairs is slightly higher when using some pairs of each of these possibilities.

Of course all methods are completely equivalent when calculating the mean. However, when applying cuts or using fits of a histograms the last approach provides a larger number of histogram entries and thus generates more robust results. On the other hand the error determination becomes slightly more complicated. Assuming that the standard deviations of the signal, σ_s, and the background, σ_b, are equal, the error formula reads

$$\Delta(S-B) = \sigma_{S-B}\sqrt{\frac{1}{2}\left(\frac{1}{n_s}+\frac{1}{n_b}\right)} \tag{7}$$

where n_s and n_b refers to the fraction of events in the fit region and σ_{S-B} is the variance of the obtained $S-B$ distribution[3]. Note that this method can make use of all available signal and background events, since their number do not need to be the same. A derivation of the error formula can be found in appendix A.3

Possible problems and concerns of Method 4:

- *Are data points multiple used?*
  Method 4 is a proper convolution of two given distributions. The error evaluation has to be performed correctly, then the statistical significance is still determined by the number of used (original) data points.
- *Is the time correlation not used?*
  Surely, short term variation will cause a wider distribution when averaging over the full run (instead of few seconds). Looking at fig. 5 this effect can be well observed for run 1343. There the width increases from 0.38 to 0.45 (Method C to Method D). On the other hand, for most of the runs this effect is small, indicating a stable conditions.
- *Does the error calculation become more complicated?*
  Yes, but only a little, to first approximation (assuming $n_s = n_b = n$ and $\sigma_s = \sigma_b$) the error is still given by

  $$\Delta(S-B) = \frac{\sigma_{S-B}}{\sqrt{n}} \tag{8}$$

  This means, the pure statistical error is not changed, but the fit becomes more robust (the effect is quantified in the appendix).

Advantage and properties of Method 4:

- A possible systematic decrease or increase in signal could be due to the observed dependence on the pairing of signal and background events. For instance, if the signal-level is systematically falling, and the measured background B is selected after the signal S is selected. The resulting difference $S-B$ will be overestimated. On the other hand, if the background B is measured before the corresponding signal, $S-B$ will be underestimated.

  A systematic combination of any pairing avoids such effects.
- This method is a correct convolution of signal and background, taking into account all available information on the distribution of signal and background. This for instance means, that a "bad" background event can not completely destroy a "good" signal event (or vice versa). Of course, this is only justified if event-to-event correlation are negligible.
- The maximal number of available background events is used, i.e. all measured background and signal events can be included in the analysis. It is not required to have *complete* pairs with equal numbers of signal and background events.

  Nevertheless, the method guaranties, that each individual background (or signal) event is used with the same share.
- Since the distribution becomes very well defined, the dependence of them mean signal on the fit methods (χ^2 or likelihood) becomes negligible.

[3]Practically the width is calculated from the FWHM with the relation $\sigma = FWHM/2.35$.

3 Combining runs of different polarity to obtain asymmetries

3.1 Error propagation

After the background subtraction a mean CsI signal X can be assigned to each run. The error ΔX is fixed by statistics and determined according to Eq. (7). For the determination of asymmetries these individual values of CsI signals for different magnet polarities are combined to pairs. Then the asymmetry is calculated from the formula

$$A = \frac{X_- - X_+}{X_- + X_+} , \tag{9}$$

where the subscript $\pm$ refers to the magnet polarity. The error on the asymmetry ΔA is than fixed by the error on the X values and can be derived by usual error propagation

$$\Delta A = \frac{2X_- X_+}{[X_- + X_+]^2} \sqrt{\left(\frac{\Delta X_-}{X_-}\right)^2 + \left(\frac{\Delta X_+}{X_+}\right)^2} . \tag{10}$$

Of course this formula only holds, if the $\Delta X \ll X$.

3.2 Combining all runs

The data sample of the June measurements can be used to build $N = 52$ independent pairs (104 runs using the W-target). A combined result can be found employing the known error estimate via

$$< A > = \frac{\sum_{i=1}^{N} \frac{1}{\Delta A_i{}^2} A_i}{\sum_{i=1}^{N} \frac{1}{\Delta A_i{}^2}} . \tag{11}$$

Correspondingly, the error is determined by

$$\Delta < A > = \left(\sum_{i=1}^{N} \frac{1}{\Delta A_i{}^2} \right)^{-\frac{1}{2}} . \tag{12}$$

The quality of this estimate might be judged by looking at the *pull distribution* and a corresponding value for χ^2,

$$\chi_i = \frac{A_i - < A >}{\Delta A_i{}^2} , \qquad \chi^2 = \sum_{i=1}^{N} \chi_i{}^2 . \tag{13}$$

4 Results on positron asymmetries

4.1 Run selection

The results are obtained by analysing 52 pairs of runs, all with W target (1226-1235, 1339-1348, 1353-1362, 1367-1378, 1385-1396, 1427-1438, 1448, 1450-1456, 1472-1481, 1496-1515). In total this corresponds to about *6 hours* of data taking. For pairing of runs only (almost) adjacent runs have been used. There is no other criterion used to reduce the number of runs.

4.2 Data selection

From each run an number of signal (undulator on) and background (undulator off) events are selected according to the following conditions

- trigger is either "beam on & undulator on" (signal) or "beam on & undulator off" (background)
- toroid "tor6130" is in the range [180 . . . 360].

The background subtracted signal is obtained with four different methods:

Method A uses adjacent signal and background events only. This corresponds to method 1 in sec. 2.2.

Method B is similar to method A but uses a different pairing of signal and background (starting with event 2). This corresponds to method 2 in sec. 2.2.

Method C averages the background by using different combinations of signal and background, but with a maximal distance of 8 events.

Method D generate a proper convolution of the signal and background distribution by using all possible combinations of signal and background events. This corresponds to method 4 in sec. 2.2.

The mean X of the corresponding distributions is determined by iteratively fitting the peak region ($\pm 2\sigma$) with a Gauss distribution[4]. The error ΔX is fixed by the width of the distribution, and the used data point (see also appendix). *Method C* is the preferred choice. Results from Methods A, B and D are determined for a comparison. [5]

Then, for each run pair an asymmetry A is calculated according to eqs. (9) and (10). The combined result of all 52 pairs is determined using eqs. (11) and (12).

4.3 Results

Table 1: Compilation of the final asymmetries with Method A-D

Method	Asymmetry	Error	χ^2/ndf
A	0.00665	0.00124	74.4/51
B	0.00900	0.00124	75.2/51
C	**0.00780**	**0.00102**	67.3/51
D	0.00743	0.00111	57.3/51

Note, that the error estimate is pure statistical; the χ^2/ndf ratio indicates that the true error is slightly higher.

5 Discussion

For the next run it is, of course, highly advisable to spend additional time on the discovery of the source of the background, and – if possible – its reduction. Nevertheless, the existing data give already a striking evidence for the existence of positron polarisation:

$$\boldsymbol{A}_{\textbf{Method C}} = \mathbf{0.0078 \pm .0010} \tag{14}$$

Note, that the error estimate is pure statistical, systematic uncertainties (like fringe field effects) remain to be determined.

Some general comments on the data analysis:

- Fitting the peak region of the $S_i - B_j$ distribution results in smaller uncertainties since it neglects the non-Gaussian tails.

[4] In case of method C and D, the sum of two Gauss distribution with a common mean is used in oder to get better fit results

[5] Note that Method D is the best (most robust) choice, if the background is stable. In the current data this is not the case, which results in a slightly higher uncertainty of the final result, see table 1.

- The new method to combine several signal and background measurement yields more robust results compared to the standard method of using only adjacent signal and background events.
- When generating secondary observables the usage of different pairings leads to more stable fit results. This also avoids possible systematic problems, and includes of all available information. The statistical error is determined by the number of (used) original data points.
- The mean asymmetries of all pairs is calculated by a weighted mean (cf. eq. (11)), i.e. without putting data into histograms. This takes all individual errors into account properly, and avoids problems from binning or fitting.
- It was noticed that fitting the data can induce additional uncertainties (cf. appenix). Using the build-in ROOT routines, the Log-likelihood method was found to give better results compared to the χ^2 method, since it has a better treatment of bins with small number of entries. For larger number of entries both approaches performed equally well.

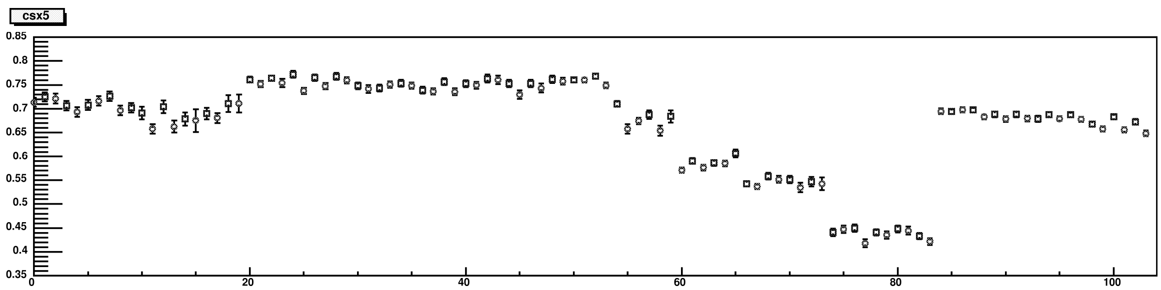

Figure 1: CsI [MeV]/Tor6130 ADC counts for 104 runs.

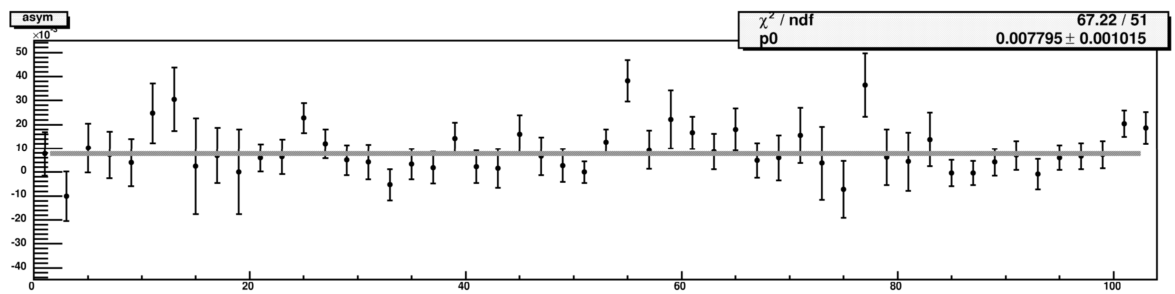

Figure 2: Asymmetries for 52 run pairs.

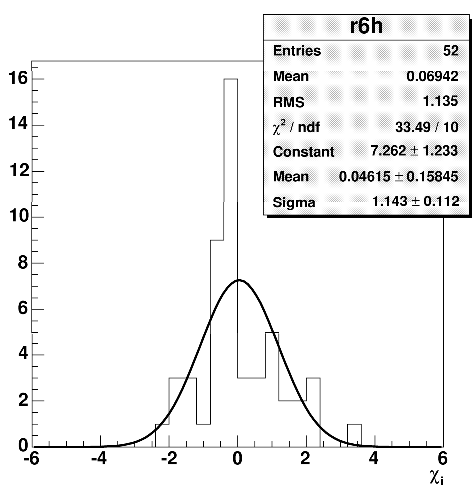

Figure 3: Pull distribution (Method C).

Table 2: CsI [MeV]/Tor6130 ADC counts for 104 runs, and resulting asymmetries.

	runno	mag	(npair)	csx5 +- csx5err	(c(-60)-c(+60))/(c(-60)+c(+60))
0	1226	60	(939)	0.713683 +-0.00901852	
1	1227	-60	(934)	0.724845 +-0.00971356	0.00775948 +-0.00920904
2	1228	60	(860)	0.721586 +- 0.0107973	
3	1229	-60	(938)	0.706532 +- 0.010036	-0.0105408 +- 0.0103148
4	1230	60	(934)	0.693567 +- 0.0098047	
5	1231	-60	(927)	0.707642 +- 0.010617	0.0100449 +- 0.0103061
6	1232	60	(939)	0.715893 +- 0.010007	
7	1233	-60	(936)	0.726671 +-0.00988802	0.00747133 +-0.00975334
8	1234	60	(943)	0.69634 +-0.00996465	
9	1235	-60	(929)	0.702199 +-0.00958056	0.00418906 +-0.00988576
10	1339	-60	(936)	0.691746 +- 0.0131547	
11	1340	60	(928)	0.65822 +- 0.0102641	0.024835 +- 0.0122887
12	1341	-60	(930)	0.703964 +- 0.012912	
13	1342	60	(938)	0.662304 +- 0.0127191	0.0304925 +- 0.0132657
14	1343	-60	(944)	0.67849 +- 0.0140537	
15	1344	60	(941)	0.67553 +- 0.0231956	0.00218634 +- 0.0200502
16	1345	-60	(940)	0.69009 +- 0.0121548	
17	1346	60	(935)	0.681027 +- 0.0103646	0.00660958 +- 0.0116383
18	1347	-60	(937)	0.710815 +- 0.017447	
19	1348	60	(938)	0.71067 +- 0.0181632	0.000102315 +- 0.0177177
20	1353	-60	(938)	0.76126 +- 0.0061586	
21	1354	60	(936)	0.751712 +-0.00615987	0.00631127 +-0.00575732
22	1355	-60	(936)	0.764002 +-0.00599786	
23	1356	60	(935)	0.753908 +-0.00910167	0.00665029 +-0.00720004
24	1357	-60	(932)	0.771916 +-0.00683722	
25	1358	60	(937)	0.737677 +-0.00667144	0.0226814 +-0.00632616
26	1359	-60	(936)	0.765233 +-0.00658078	
27	1360	60	(932)	0.747089 +-0.00646076	0.0119974 +- 0.0060971
28	1361	-60	(935)	0.767304 +-0.00699228	
29	1362	60	(941)	0.75963 +-0.00642792	0.00502579 +-0.00621769
30	1367	-60	(934)	0.748448 +-0.00726434	
31	1368	60	(928)	0.741545 +-0.00800396	0.00463298 +-0.00725771
32	1369	-60	(906)	0.743271 +-0.00720087	
33	1370	60	(928)	0.751116 +-0.00659396	-0.00524943 +-0.00653678
34	1371	-60	(938)	0.753362 +-0.00699795	
35	1372	60	(935)	0.748226 +-0.00658742	0.00342024 +-0.00639907
36	1373	-60	(925)	0.738921 +-0.00733553	
37	1374	60	(935)	0.736165 +-0.00663184	0.00186817 +-0.00670273
38	1375	-60	(927)	0.756411 +-0.00691057	
39	1376	60	(929)	0.735438 +-0.00725559	0.0140584 +-0.00672173
40	1377	-60	(928)	0.752694 +- 0.007101	
41	1378	60	(913)	0.749248 +-0.00747799	0.00229392 +-0.00686684
42	1385	-60	(925)	0.762818 +-0.00823714	
43	1386	60	(921)	0.760462 +-0.00909076	0.00154651 +- 0.0080546
44	1387	-60	(927)	0.752857 +-0.00811831	
45	1388	60	(908)	0.729192 +-0.00872903	0.0159677 +-0.00805371
46	1389	-60	(936)	0.753092 +-0.00779965	
47	1390	60	(931)	0.743306 +-0.00898905	0.00653921 +- 0.0079607
48	1391	-60	(934)	0.761696 +-0.00769676	
49	1392	60	(935)	0.757334 +-0.00712403	0.00287136 +-0.00690271
50	1393	-60	(930)	0.760126 +-0.00506664	
51	1394	60	(938)	0.760183 +-0.00464749	-3.80214e-05 +-0.00452233
52	1395	-60	(931)	0.768347 +-0.00525365	
53	1396	60	(935)	0.748941 +-0.00575474	0.0127901 +-0.00514198

Table 3: Table 1 continued

	runno	mag	(npair)	csx5 +- csx5err	(c(-60)-c(+60))/(c(-60)+c(+60))
54	1427	-60	923	0.710197 +-0.00591859	
55	1428	60	928	0.65812 +-0.00986468	0.0380591 +-0.00856263
56	1429	60	933	0.674837 +-0.00649533	
57	1430	-60	896	0.687667 +-0.00874769	0.00941674 +- 0.0079752
58	1431	60	923	0.654514 +- 0.0104047	
59	1432	-60	927	0.683961 +- 0.0124871	0.0219999 +- 0.0120981
60	1433	60	1376	0.570748 +-0.00517849	
61	1434	-60	928	0.590141 +-0.00575751	0.016706 +-0.00665969
62	1435	60	933	0.576251 +-0.00633145	
63	1436	-60	931	0.586168 +-0.00600151	0.00853123 +- 0.0075086
64	1437	60	929	0.58495 +-0.00639608	
65	1438	-60	912	0.606516 +-0.00838491	0.0181008 +-0.00881023
66	1448	-60	910	0.541969 +-0.00550463	
67	1450	60	913	0.536774 +-0.00557088	0.0048161 +-0.00726053
68	1451	-60	915	0.558159 +-0.00723461	
69	1452	60	916	0.551739 +-0.00750635	0.00578419 +-0.00939511
70	1453	-60	924	0.551006 +- 0.0075873	
71	1454	60	909	0.534315 +-0.00993549	0.015379 +- 0.0115664
72	1455	-60	922	0.546508 +-0.00997346	
73	1456	60	924	0.541859 +- 0.013369	0.00427181 +- 0.0153439
74	1472	-60	919	0.441064 +-0.00724966	
75	1473	60	920	0.447263 +-0.00770959	-0.00697854 +- 0.0119084
76	1474	-60	926	0.449915 +-0.00785492	
77	1475	60	912	0.418492 +-0.00825268	0.0361843 +- 0.0131517
78	1476	-60	916	0.441029 +-0.00689138	
79	1477	60	933	0.43551 +-0.00749562	0.00629624 +- 0.0116226
80	1478	-60	905	0.448056 +-0.00724733	
81	1479	60	899	0.444411 +-0.00817477	0.00408429 +- 0.0122472
82	1480	-60	892	0.433654 +-0.00651755	
83	1481	60	911	0.421676 +-0.00705572	0.0140032 +- 0.0112435
84	1496	60	942	0.694713 +-0.00588817	
85	1497	-60	927	0.694176 +-0.00490865	-0.000387023 +-0.00551904
86	1498	60	935	0.698106 +-0.00529286	
87	1499	-60	934	0.697434 +- 0.0047302	-0.000481492 +-0.00508631
88	1500	60	919	0.683009 +-0.00479175	
89	1501	-60	933	0.688731 +- 0.0060453	0.00417153 +-0.00561824
90	1502	60	933	0.678738 +-0.00605843	
91	1503	-60	929	0.688526 +-0.00574226	0.00715903 +-0.00610764
92	1504	60	937	0.67998 +-0.00592342	
93	1505	-60	926	0.678957 +-0.00630677	-0.000752633 +-0.00636726
94	1506	-60	1549	0.687842 +-0.00505909	
95	1507	60	1867	0.679571 +-0.00475708	0.00604874 +-0.00507667
96	1508	-60	1876	0.68748 +-0.00486474	
97	1509	60	1875	0.678292 +-0.00561821	0.00672736 +-0.00544674
98	1510	-60	1866	0.667386 +-0.00561928	
99	1511	60	1871	0.657846 +-0.00520153	0.00719865 +-0.00577493
100	1512	-60	1875	0.683113 +-0.00501437	
101	1513	60	1876	0.656038 +-0.00530614	0.0202174 +-0.00545901
102	1514	-60	1867	0.67285 +-0.00604692	
103	1515	60	1858	0.648489 +-0.00619094	0.0184362 +-0.00655343

chisqr= 67.2579 ndf= 51 0.00780441 +-0.00101492

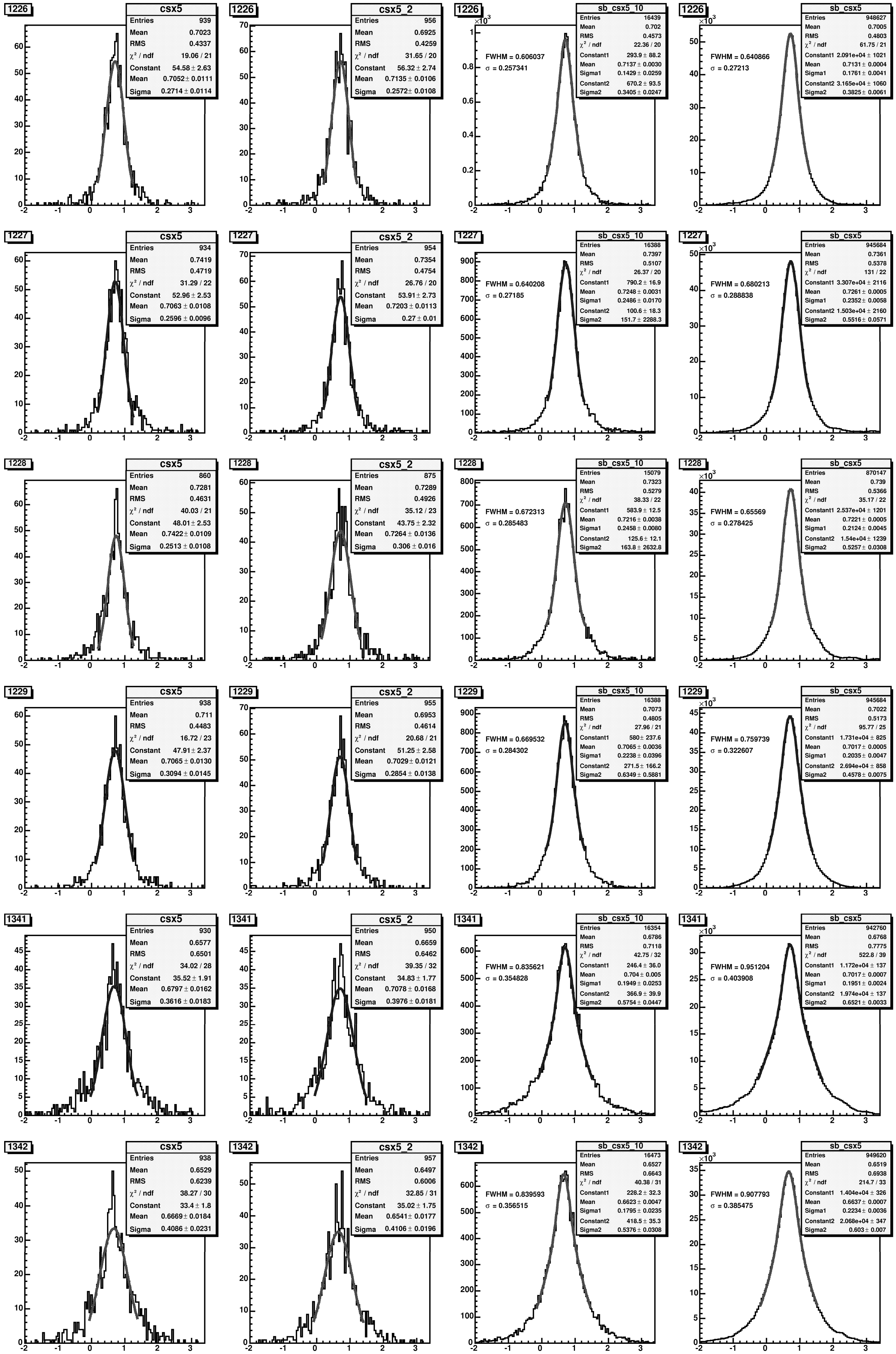

Figure 4: Fit background subtracted CsI signal (central crystal) for selected runs using methods A - D (from left to right).

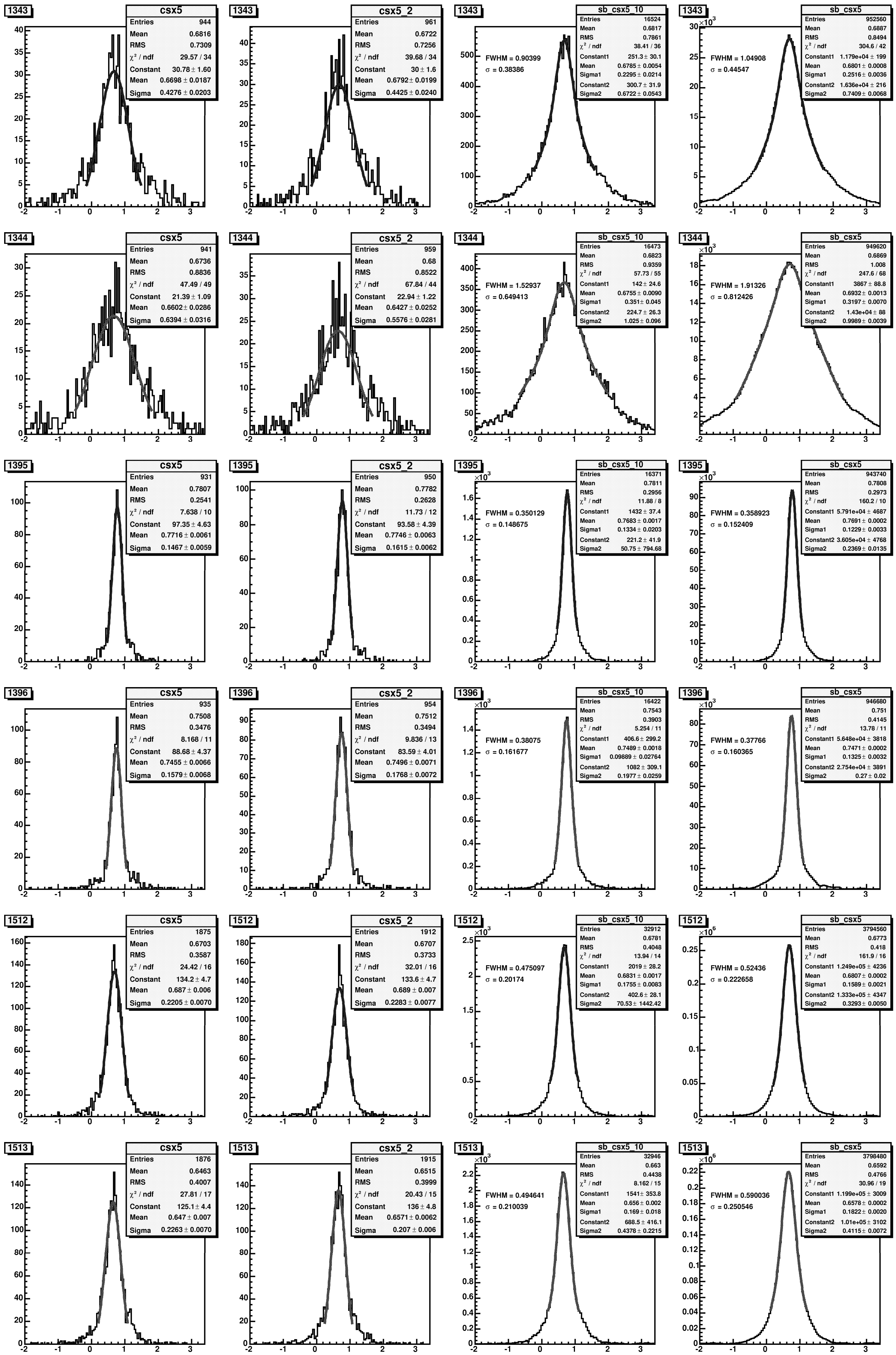

Figure 5: Fit background subtracted CsI signal (central crystal) for selected runs using methods A - D (from left to right).

A Appendix: Toy Monte Carlo

To illustrate the features of the proposed method, and to validate the error estimate, a toy Monte Carlo is employed.

A.1 Method

1. Generate $n = 1000$ signal and background events according to a Gauss distribution

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \frac{(x-\mu)^2}{2\sigma^2} \tag{15}$$

with $\mu = 0$ and $\sigma = 1$. In this context signal idealises the "undulator on" events, and background the "undulator off" events, which is, of course, a simplification.

2. Determine the estimate of the mean of the "signal-minus-background" by subtracting the means of all signal and background events.

$$\mu_{\text{means}} = \sum_{i=1}^{n} s_i - \sum_{i=1}^{n} b_i \tag{16}$$

This estimate should vary around the true value with a variance of

$$\Delta\mu_{\text{means}} = \frac{\sigma}{\sqrt{2n}} \approx 0.04472 \tag{17}$$

3. Create a distribution of signal minus background in three different ways:
   - Plot all pairs $(s_i - b_i)$ for $i = 1, \ldots, n$.
   - Plot all pairs $(s_i - b_{i+1})$ for $i = 1, \ldots, (n-1)$ and the pair $(s_n - b_1)$.
   - Plot all pairs $(s_i - b_j)$ for $i, j = 1, \ldots, n$.

All distributions have the same mean value μ_{means} (since they use the same data). In the limit $n \to \infty$ all these distributions are equivalent and should resemble a Gauss distribution with a width of $\sigma_{s-b} = \sqrt{2}\sigma$.

4. Fit the peak region $[-1.5\,\sigma_{s-b}, 1.5\,\sigma_{s-b}]$ of all distributions. Thus only 86.6% of the data points are used in the determination of the mean, leading to a 7.43% increase in the uncertainty

$$\Delta\mu_{\text{fit}} = \frac{1}{\sqrt{.866386}} \Delta\mu_{\text{means}} \approx .04805 \tag{18}$$

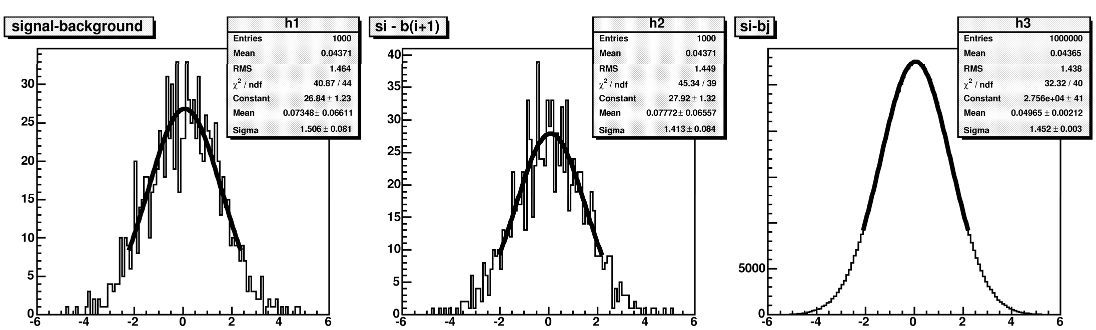

Figure 6: Analysis of a single Monte Carlo sample for 1000 signal minus 1000 background events. Both signal and background events are generated according to Gauss distributions ($\mu = 0$, $\sigma = 1$). Plotted are two ways to produce "adjacent" pairs (left and central plot), as well as the distribution of all possible signal background combination (right plot). The fits are generated by iteratively fitting a $\pm 1.5\sigma$ peak region.

5. Now repeat step 1. to 4. $N = 5000$ times in order to get distributions of the fluctuations.

A.2 Results

Figure 7 shows the obtained distributions of fit values for the different methods, while figure 8 gives the distribution if using the exact formula, i.e. all data points not only the peak region. These plots clearly show, that doing a complete convolution of signal and background events gives more stable results. The variation of the exact means agrees very well with the expect width, cf. Eq. (17). Also the variation of the full convolution (left plot in figure 7) matches the estimate in Eq. (18). A summary of obtained results can be found in Tab. 4.

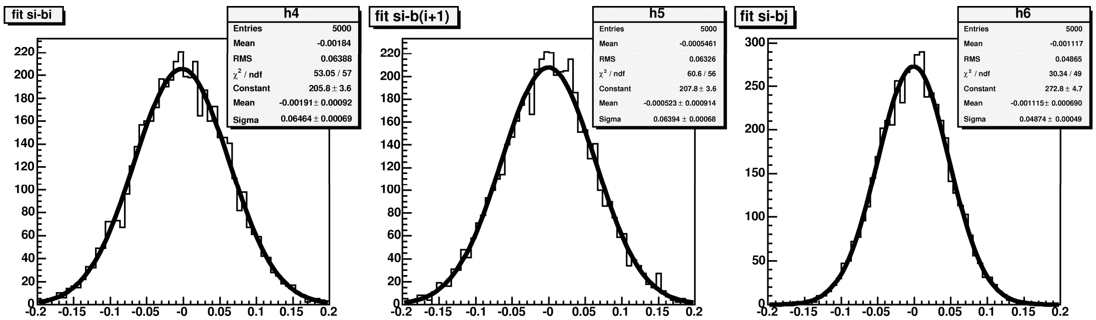

Figure 7: Distribution of variations of fit mean using the different methods. The left and central plot give the distribution when using only "adjacent" pairs, while the right plot shows the variation of fit means of the full convolution.

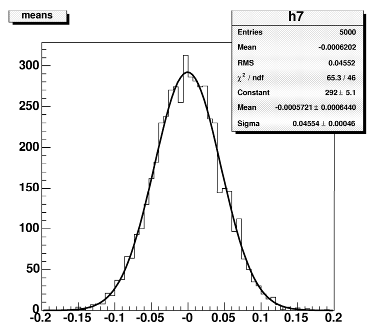

Figure 8: Distribution of variations using the exact mean using Eq. (16).

fit region	data fraction (ideal Gauss)	mean of all data $s_i - b_j$	mean of all data calculated	"adjacent" points $(s_i - b_i)$	"adjacent" points $(s_i - b_{i+1})$	the full convolution simulated	the full convolution calculated
$[-\sigma, \sigma]$	0.682689	.0455(5)	.04472	.0914(13)	.0914(13)	.0518(5)	.0541
$[-1.5\sigma, 1.5\sigma]$	0.866386	-	-	.0646(7)	.0639(7)	.0487(5)	.04805
$[-2\sigma, 2\sigma]$	0.954500	-	-	.0525(5)	.0526(5)	.0467(5)	.04577
$[-3\sigma, 3\sigma]$	0.997300	-	-	.0463(4)	.0464(5)	.0456(5)	.04478
$[-6, 6]$	0.999978	-	-	.0455(5)	.0456(5)	.0455(5)	.04472

Table 4: Variation of fit results as determined by the toy Monte Carlo. The fit range is varied from $\pm 1\sigma = \sqrt{2}$ up to $\pm 4.25\sigma = 6$.

One can conclude, that using the distribution after the full convolution of signal and background events yields more stable results. Note that the fit uncertainties of "adjacent" fits become even larger if the χ^2 fit routines are used instead of Poisson-log-likelihood, underlining the fact that the likelihood method is more appropriate for this problem. On the other hand the impact of the fit method to the $(s_i - b_j)$ sample is negligible.

A.3 Error formula

Here, a derivation of the error formula Eq. (7) used in Sec. 2.2 shall be given.

Assume a signal distribution (including some background contributions) S_i and (pure) background distribution B_i. Both distributions are assumed to be dominated by background, and thus have the same variance σ. A background subtracted signal $S - B$ can now be obtained from n_s signal and n_b background measurements,

$$\begin{aligned} < S > &= \frac{1}{n_s} \sum_{i=1}^{n_s} s_i \quad \text{and} \\ < B > &= \frac{1}{n_b} \sum_{i=1}^{n_b} b_i \ , \end{aligned} \tag{19}$$

by just taking the difference:

$$< S - B > = < S > - < B > \ . \tag{20}$$

Then, the error of the mean background subtracted signal $< S - B >$ is determined by the individual errors on $< S >$ and $< B >$, which can be estimated by

$$\Delta < S > = \frac{\sigma}{\sqrt{n_s}} \qquad \text{and} \qquad \Delta < B > = \frac{\sigma}{\sqrt{n_b}} \ . \tag{21}$$

Then one obtains

$$\begin{aligned} \Delta < S - B > &= \sqrt{(\Delta < S >)^2 + (\Delta < B >)^2} \\ &= \sigma \sqrt{\frac{1}{n_s} + \frac{1}{n_b}} \ , \end{aligned} \tag{22}$$

which can be related to the width of the distribution $(S_i - B_i)$ via ${\sigma_{S-B}}^2 = {\sigma_S}^2 + {\sigma_B}^2 = 2\sigma^2$ resulting in

$$\Delta < S - B > = \frac{\sigma_{S-B}}{\sqrt{2}} \sqrt{\frac{1}{n_s} + \frac{1}{n_b}} \ . \tag{23}$$

An independent confirmation of this error formula was already obtained in section A.2 via Monte Carlo techniques. Note, that in the case where a limited number of data points is used for the evaluation of the mean (e.g. by fitting only the peak region of the distribution), the numbers of signal and background points, n_s and n_b, have to be reduced correspondingly.

A.4 Quality measures

So far, no effort has been made to select "good" or "bad" runs, i.e. all available data has been used. Possible criteria are

- correlation of signal (or background) to toroid signal (note that this is particular visible in runs where the toroid varies strongly!),
- correlation of signal+background to background,
- width of the background subtracted signal distribution, $(S_i - B_j)$,
- the change of the width of the $(S_i - B_j)$ distribution when moving from "16 event window" (Method C) to the full convolution (Method D),
- χ^2/ndf of the fitted $(S_i - B_j)$ distribution.

An online analysis, i.e. immediately after a run, might help to observe problems in the data taking.