Background Subtraction in the R-scan

Derrick Kong
University of Hawaii

June 7th, 2000

Abstract

I review background subtraction methods for removing beam-gas/beam-wall events from the R-scan data. Results from the background subtraction scheme used for the 1998 R-scan conducted at BES/BEPC are compared against two newer schemes involving vertex fitting. Overall results for the schemes are generally consistent with each other. A simple averaging scheme for vertex fitting is recommended for use in the 1999 R-scan.

1 Introduction

The experimental determination of R, the ratio of $e^+e^- \to$ hadrons divided by $e^+e^- \to \mu^+\mu^-$, is calculated via the following formula,

$$R = \frac{N_{had} - N_{back}}{\sigma^0_{\mu\mu} \cdot \mathcal{L} \cdot \epsilon_{had} \cdot \epsilon_{trg} \cdot (1+\delta)}. \tag{1}$$

The factor of interest for this work is N_{back}, the number of background events. This factor can be further divided as follows,

$$N_{back} = N_{beamgas} + \Sigma_l N_{ll} + N_{\gamma\gamma}, \tag{2}$$

where $N_{beamgas}$ is the contribution from beam-gas/beam-wall events, $\Sigma_l N_{ll}$ is the contribution from misidentified lepton pairs, and $N_{\gamma\gamma}$ is the contribution from 2 gamma events. The lepton and photon pairs are easily removed via statistical subtraction, but $N_{beamgas}$ is the most serious source of background in terms of numbers of events and difficulty in removal.

For the 1998 R-scan, separated and single beam data were used to calculate the beam-gas contribution. For the 1999 data, some beam-gas events were explicitly filtered in the hadronic selection process, but the rest were subtracted out by fitting the z-vertex of the events. Below, I review the beam-gas method and describe two methods of determining the z-vertex for fitting.

2 98 R-scan Data

For the 98 R-scan data, following method was used:

$$N_{beamgas} = f \times N_{sep}, \tag{3}$$

$$f = \frac{\int_{run} P_{run} \cdot I_{run}\ dt}{\int_{sep} P_{sep} \cdot I_{sep}\ dt}, \tag{4}$$

where N_{sep} is the of events from separated beam data passing the hadronic selection routine, and the factor, f, is defined by the ratio of the integrated gas pressure, P, and beam current, I, for the colliding beam data (run) and separated beam data (sep). [1] This method was used by Crystal Ball for their R-scan. [2]

In the 98 R-scan, $N_{beamgas}$ ranged anywhere from $\approx 2-20\%$ of N^{obs}_{had}. Poor data quality at 5.0 GeV lead to an especially large beam-gas contribution. See Figure 1 for a plot of ratio of beam-gas to hadronic events.

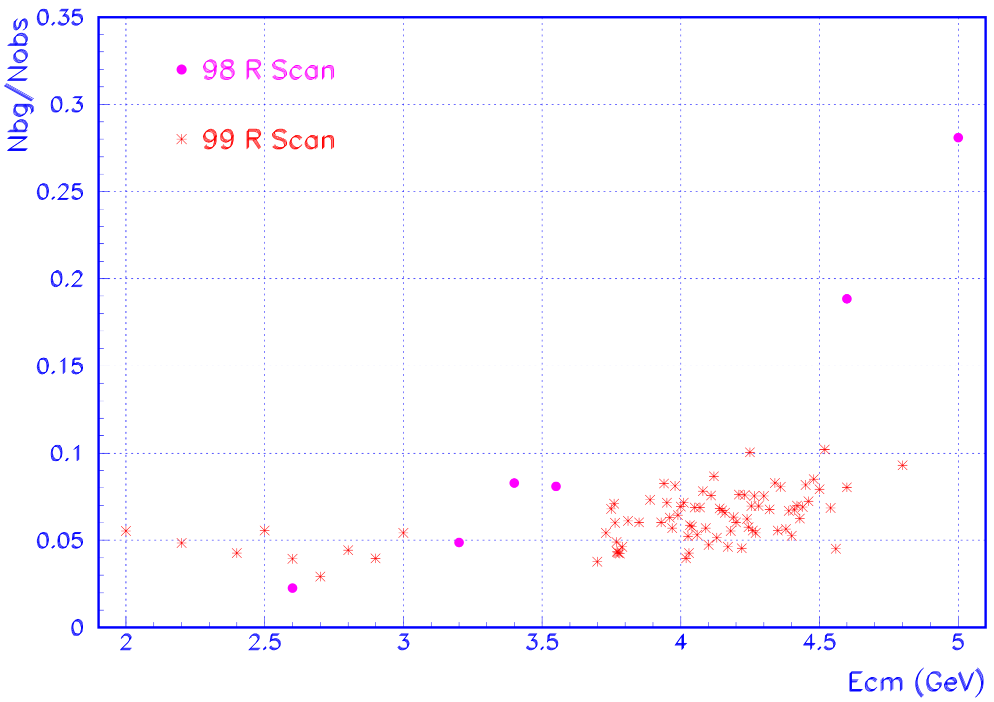

Figure 1: Plot of the ratio of beam-gas events to hadronic-events at each energy point. Figure from Huang Guangshun's talk at BES00.

3 Vertex Determination via Simple Averaging

Colliding beam events are expected to have a Gaussian distribution in z (coordinate along the beam line) around the interaction point (IP). On the other hand, beam-gas events are expected to have a relatively flat distribution in z. Thus, it is expected that a fit of the z-vertex of events to a Gaussian plus background will allow a calculation of the true number of colliding beam events.

As vertexing info from the BES vertex chamber is not available, a simple definition of the vertex is instead used,

$$V_{z,evt} = \frac{\Sigma_{i=1}^{N} V_{zi,mfit=2}}{N}, \tag{5}$$

where $V_{z,evt}$ is the z-coordinate of the event vertex, N is the number of tracks in the event, and V_{zi} is the z-coordinate of the i'th track at its closest approach to the IP. Only MFIT=2 (good helix fit) tracks are used to calculate this average (Note that the hadronic selection routine requires each good hadronic event to have at least one MFIT=2 track). More details of this method are available in an analysis memo. [3]

See Figures 2 and 3 for examples of the Gaussian fit to the z-vertex distribution of this method.

4 Vertex Determination via Fitting Routine

In addition to the simple averaging method for vertex determination, a more sophisticated method was investigated. Using a routine in BES library, `vxkwikj.f`, which was taken from Mark III (and before that from Mark II), fits were performed to the same data set as a cross-check.

The fitting routine attempts to find intersection points among all good (MFIT=2) tracks and then finds a least-squares central point between all the intersections. The fit requires ≥ 2 MFIT=2 tracks and at least one pair of non-parallel tracks. Because of this requirement, the routine is unable to find a good vertex in approx 1% of hadronic events (as a "good" hadronic event is only required to have at least 1 MFIT=2 track). In such cases, in order to have comparable results, the routine falls back to using the simple averaging method of the previous section to determine a vertex.

The routine also has an option to select the fidicial volume for usable tracks. A tight cut uses requires that the closest approach of the track to the IP have $r < 1.5$ cm and $z < 15$ cm, while a loose cut only requires the closest approach have $r < 15$ cm. Both tight and loose fiducial cuts were used to determine which was best. Additional details are available in an analysis memo. [4]

See Figures 2 and 3 for examples of the Gaussian fit to the z-vertex distribution of this method.

5 Numerical Comparisons

Here I present numerical comparisons of the above analysis methods. In order to facilitate comparisons, I present analysis of the 1998 data points, where the separated-beam method results are already available. Analysis of 1999 R-scan data gives comparable results. [4]

The first comparison made is the net number of hadrons (candidate hadrons minus beam-gas contribution) obtained via each method. As seen in Table 1, the averaging method

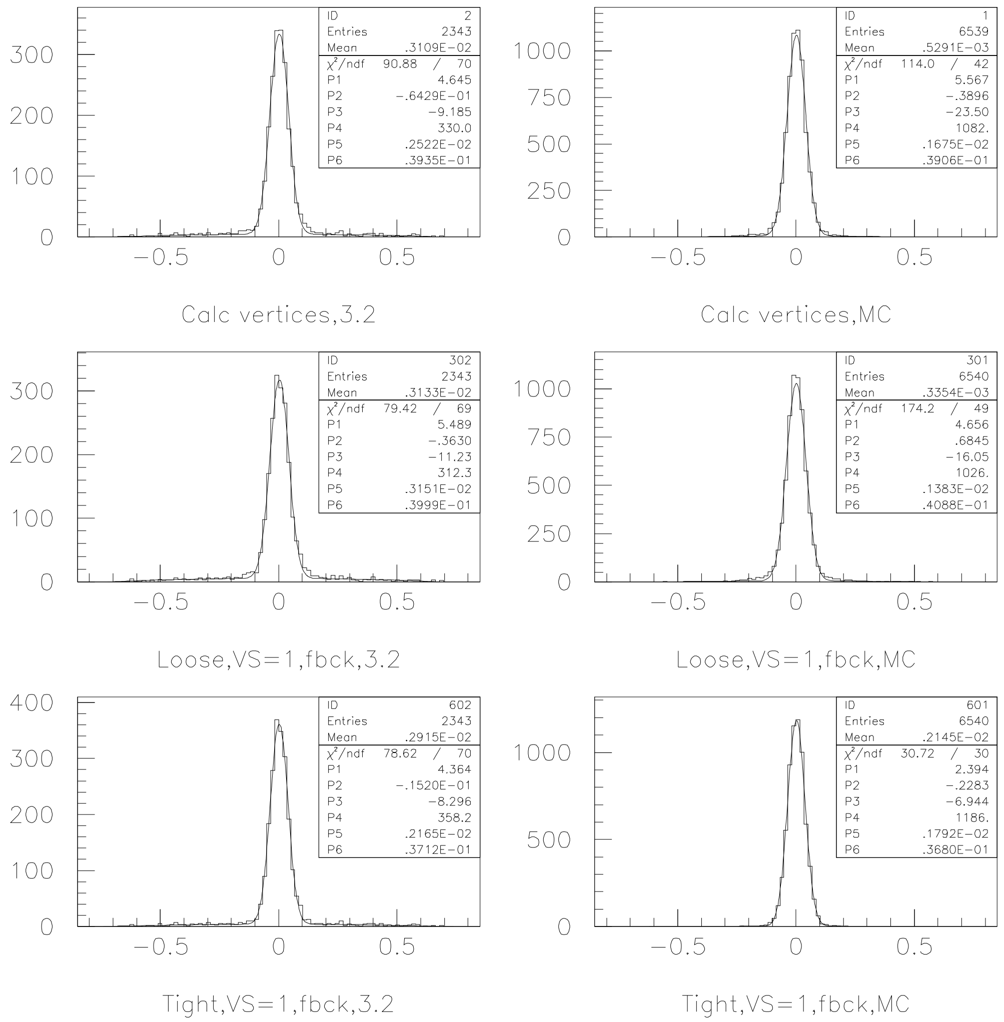

Figure 2: Gaussian plus 2nd order polynomial fits to vertex distributions for 3.2 GeV data. Left hand side is data, right hand side is Monte Carlo. Upper row is the simple averaging method, the middle row is vertex fitting with loose fiducial cuts and the bottom row is vertex fitting with tight fiducial cuts. This figure is typical of most data points.

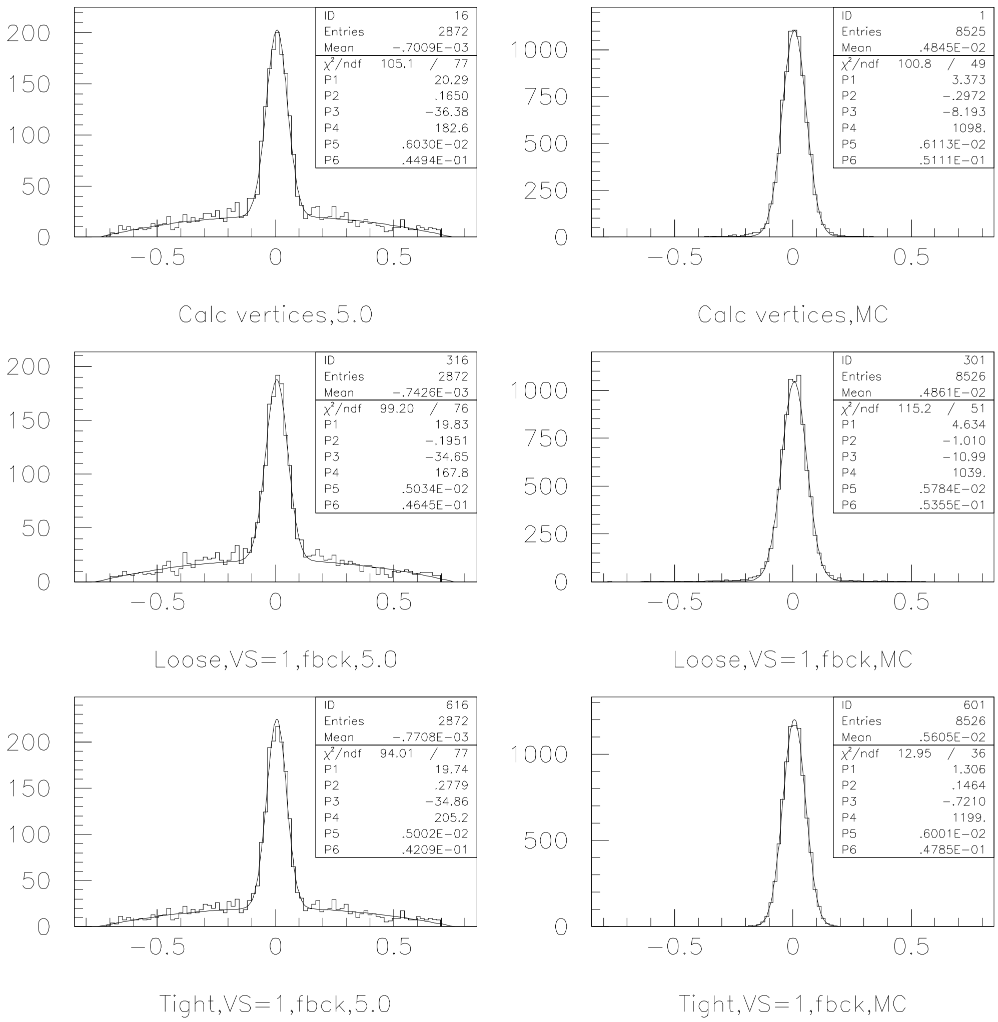

Figure 3: Gaussian plus 2nd order polynomial fits to vertex distributions for 5.0 GeV data. Left hand side is data, right hand side is Monte Carlo. Upper row is the simple averaging method, the middle row is vertex fitting with loose fiducial cuts and the bottom row is vertex fitting with tight fiducial cuts. This figure shows the large contribution from beam-gas.

and tight cut fit method are consistent to within a few percent, while the loose cut fit is generally too low. Note the discrepancy at 5.0 GeV; this is attributed to the high beam-gas contamination due to poor data quality.

E_{cm}	Sep beam	Avg method		Fit method, loose cut		Fit method, tight cut	
(GeV)	# had	# had	% diff	# had	% diff	# had	% diff
2.60	5490	5257	-4.2	5133	-6.5	5413	-1.4
3.20	1951	1907	-2.3	1841	-5.6	1961	+0.5
3.40	1971	1906	-3.3	1818	-7.8	1965	-0.3
3.55	2457	2381	-3.1	2311	-5.9	2496	+1.6
4.60	1215	1165	-4.1	1113	-8.4	1211	-0.3
5.00	1185	1214	+2.5	1149	-3.0	1273	+7.4

Table 1: Comparison of net number of hadrons ($N_{had} - N_{beamgas}$) obtained by different methods.

In addition to the number of hadrons, the hadronic detection efficiency needed to be checked. The various routines were run on Monte-Carlo samples of 10,000 events generated via the JETSET generator and efficiencies relative to the separated beam method are summarized in Table 2. Again the averaging method and tight cut fit method agreed reasonably well, with the loose cut fit being somewhat lower in efficiency.[1]

E_{cm}	Change in Efficiency (%)		
(GeV)	Avg method	Fit method, loose cut	Fit method, tight cut
3.20	-4.5	-3.8	+0.1
3.40	-2.7	-2.9	-0.4
3.55	-2.1	-3.6	+0.7
4.60	-1.4	-2.6	+0.3
5.00	-2.0	-3.0	+0.1

Table 2: Efficiency of vertex fits relative to separated beam method.

The overall changes in R value relative to the beam-gas method are summarized in Table 3. All methods generally agreed to within about 5% other than at 5.0 GeV, but the averaging method and tight cut fit again were somewhat closer to the separated beam method.

Finally, as an objective evaluation of the vertex determination routines, the χ^2/dof values for the fits to the vertex distributions were tabulated in Table 4. The fit was performed using a Gaussian and 2nd order polynomial. From this one can see confirmation of the above results: the averaging and tight cut fit methods obtained better fits.

[1] The 2.6 GeV enery point is not used here and in later tables as the MC file was accidentally deleted and not available.

E_{cm}	Change in R Value (%)		
(GeV)	Avg method	Fit method, loose cut	Fit method, tight cut
3.20	+2.2	-1.8	-1.5
3.40	-0.6	-4.9	+0.9
3.55	-1.0	-2.3	-1.0
4.60	-2.7	-5.8	+0.0
5.00	+4.5	+0.0	+7.3

Table 3: Change in R value relative to separated beam method.

E_{cm}	χ^2/dof		
(GeV)	Avg method	Fit method, loose cut	Fit method, tight cut
2.60	1.382	2.066	1.076
3.20	1.298	1.151	1.123
3.40	1.712	1.290	1.566
3.55	1.499	1.951	1.239
4.60	0.991	1.429	0.840
5.00	1.365	1.305	1.221
Unwtd avg:	1.375	1.532	1.178

Table 4: Comparison of χ^2 of vertex fits.

6 Conclusions

Both the averaging method for vertex determination and tight cut vertex fitting routine are generally consistent in R value with the separated beam method to $\approx 2\%$ (except for the 5.0 GeV data point) and have reasonable χ^2 values for the Gaussian plus polynomial fit. The loose cut vertex fit has poorer consistency and a higher χ^2. However, the fiducial volume cut (which is not required or used in the averaging method) restricts the number of tracks used and may bias the result.

Thus, with no compelling preference, simplicity argues for using the averaging method for vertex determination and factoring the other methods into the systematic error.

References

[1] Bai, J.Z., *et al.* (BES collaboration), Phys. Rev. Lett. 84, 4 (2000).

[2] Edwards, C., *et al.*, SLAC-PUB-5160 (1990).

[3] Huang, G., "The Subtraction of Beam-Associated Background in Hadronic Event Sample," BES analysis memo, January 11, 2000.

[4] Kong, D., "Using a Vertex Fitting Routine for Beam-Associated Background Removal," BES analysis memo, March 19, 2000.