Procedure for Background Subtraction Underneath the Signal Region

Cedric M. Guss

12 October, 1995

Abstract

We report here the method for the subtraction of background underneath the elastic signal at each t-bin using only information from the elastic detector combination under analysis. It is found for the low-t runs 745+746 (with no main ring) that the total ratio of signal to background remaining after all analysis cuts was slightly better than 1 to 2 with about 2300 elastics available (within $\pm 3\sigma$ of the left/right xx and yy difference distribution) for physics analysis. The 1/1 point (signal to background) is reached just above t = 0.005 $(GeV/c)^2$.

I. The Background Function

In order to guard against correlations in x of y background distributions and vice versa, plots of x left side (x_l) versus x right side (x_r) were broken up into 2 mm strips along each axis, and the y distributions were plot and fit to a function of 2 Gaussians summed together; and the same was done with the x and y roles reversed to get the x distributions. Physically, the Gaussians represent the halo and beam gas contributions to the background. Thus there were 6 fit parameters (2 means, 2 sigmas, and 2 normalizations) for each strip of background data. Since the halo was sharply concentrated about the beam at the low-t region of the detectors, the function only fit the beam gas contribution above a 3 mm radius from the low-t point of the detectors. (The only exception to this rule was in fitting the y_r distributions to the x_r strips where only ONE Gaussian was used for the entire function.) Hence there were 5 x_l strips of y_l distributions, 5 x_r strips of y_r distributions, 4 y_l strips of x_l distributions, and 4 y_r strips of x_r distributions. The 4 strips in one case and 5 in another is arbitrary except that it was sensible to break the x background strips into an odd number about x=0.

In order to guarantee that the data was all background, the x strips had the additional requirement that the absolute value of (x_l - 1.45 * x_r) was greater than 25 vdas cells (or 1.125 mm), where 1.45 accounts for the ratio of left side to right side x effective lengths; this cuts out the xx diagonal elastic signal. The y strips had a similar cut, with the effective length ratio approximately equal to 1, the absolute value of (y_l - y_r) needing to be larger than 20 vdas cells (or 0.9 mm).

Table 1 lists the parameter results from the 2-summed Gaussian fit of each x and y distribution for each y and x strip in detector 1 (left side) and detector 4 (right side), respectively, for no main ring runs 745+746, with only the requirement on the data that a cluster be found on the side of interest OUTSIDE the yy and xx diagonal region and the hardware trigger be elastic (no in-time ring counters and at least one trigger counter in-time from detectors 1 and 4 with nothing in 2 or 3, and additionally the heel adc above threshold in 1 and 4):

strip in vdas units:	beam gas norm	beam gas mean	beam gas sigma	halo norm	halo mean	halo sigma
50< y_l <100	26.52	241.61	99.37	55.87	250.97	18.59
100< y_l <150	37.15	247.95	83.96	32.62	250.11	24.19
150< y_l <200	30.04	242.71	76.36	61.43	251.50	23.94
200< y_l <220	28.63	231.49	64.09	92.01	251.45	19.03
120< x_l <170	112958.84	722.72	207.37	7048199.50	447.27	48.67
170< x_l <220	375020.59	887.53	245.53	40697.99	360.77	61.25
220< x_l <270	264975.48	845.57	231.24	72469.45	329.15	47.96
270< x_l <320	1540359.00	1015.90	268.37	49.94	208.72	12.91
320< x_l <370	10841565.00	1329.50	339.74	-1304.03	256.82	23.87
50< y_r <100	21.83	249.35	73.37	—	—	—
100< y_r <150	22.67	244.68	67.99	—	—	—
150< y_r <180	19.52	242.43	62.41	11.11	243.30	24.96
180< y_r <200	15.82	239.57	60.24	17.87	240.25	22.33
110< x_r <160	3218.07	579.06	225.26	—	—	—
160< x_r <210	6926.00	595.21	203.14	—	—	—
210< x_r <260	55840.47	608.12	174.13	—	—	—
260< x_r <310	14547.70	650.12	212.94	—	—	—
310< x_r <360	173.06	272.19	120.08	—	—	—.

Table 1. Parameters for the 2-summed Gaussian background fits to the data of runs 745+746, 1-4 combination, along x and y for each y and x strip, respectively.

The slightly different divisions of one side versus the other in column 1 simply have to do with the position of the data in each detector in vdas units. Now, one can notice how the distribution sigmas tend to agree with each other from one strip to the next, giving credence to the theory that the background SHAPES do not have a correlation trend in x or y. However, as seen from the normalizations, the relative amount of one background to the other (halo to beam gas) DOES change from strip to strip, as the halo increases dramatically at low y (180-220 in vdas units) and the beam gas stays relatively constant. It isn't easy to see this effect in strips of x however, as the beam gas Gaussian fit seems to "steal some of the thunder" from the halo Gaussian fit, but the difference between the two still shows up in the sigmas where the halo Gaussian sigma is much sharper by a factor of 5 or so.

Finally, we end up with 4 separate functions for the x_l, x_r, y_l, and y_r distributions, using 2-summed Gaussians for each and the parameters given above, the parameter set determined by the x and y regions within which to the find the background. THE FINAL BACKGROUND FUNCTION IS THEN THE PRODUCT OF THESE 4 SEPARATE FUNCTIONS; and any x and y correlations are accounted for by using the x and y distribution parameters above corresponding to the y and x strips within which the background is being determined.

II. Normalization with the Data

The background function was integrated over 4-dimensional space (x and y on each side) in order to compare predicted background with observed background outside the diagonal regions. Again, plots of x_l versus x_r and y_l versus y_r (Figures 1) were made using the background function, and compared against the corresponding plots of real data. Background plots of y_l versus x_l and y_r versus x_r can also be seen in Figures 2.

The data plots had ALL analysis cuts applied EXCEPT the xx and yy diagonal region cuts, of course, in order to see the background outside the signal region. This means that adc and tdc criteria for elastics had to be met with clusters on both sides, and the cluster with maximum vdas grey count was chosen for each side. Each plot is in units of vdas cells, left side detector versus right side detector. The comparison was done by cutting out the diagonal region (defined above) from the predicted and data plots, and breaking up each plot into strips along each axis, the same ones used for finding the background fits in the first place. Then a strip by strip comparison of predicted background with real background data could be done.

The ratio of the 2-summed backgrounds (data versus predicted) within the strip was used to renormalize the fit parameter normalizations for that strip, and done in each case (x_l,x_r,y_l,y_r), so that the background function could be resummed and the strip comparisons done again. This was a several step repetitive process, and in practice, the x_l and x_r strip comparisons and renormalizations were done together in one cycle, then the y_l and y_r strip comparisons and renormalizations were done in the next cycle, and so forth. After maybe 10 such reiterations, the total background function was able to predict the real background data within 2 % for the x strip backgrounds and within 3 % for the y strip backgrounds. Table 2 shows predicted versus real background comparisons after diagonal cuts were made and iterations of background normalizations were completed:

vdas strip:	real background	predicted background
$50< y_l <100$	470	457
$100< y_l <150$	1659	1616
$150< y_l <200$	3407	3318
$200< y_l <220$	7027	6843
$120< x_l <170$	927	919
$170< x_l <220$	2704	2680
$220< x_l <270$	7209	7251
$270< x_l <320$	2316	2314
$320< x_l <370$	704	698
$50< y_r <100$	1120	1156
$100< y_r <150$	3609	3725
$150< y_r <180$	4558	4704
$180< y_r <200$	3395	3507
$110< x_r <160$	742	739
$160< x_r <210$	2817	2810
$210< x_r <260$	5409	5422
$260< x_r <310$	3913	3908
$310< x_r <360$	995	994.

Table 2. Comparisons of real versus predicted backgrounds along x and y in each detector for no main ring runs 745+746, 1-4 combination.

The background summed over each dimension should ideally be equal, but because a few different lines were purposely and arbitrarily missed along each dimension, they don't exactly agree. However, for $\sum y_l$ we get 12563 background events versus 12234 predicted, for $\sum y_r$ we get 12682 background events versus 13092 predicted, for $\sum x_l$ we get 13860 background events versus 13862 predicted, and for $\sum x_r$ we get 13876 background events versus 13873 predicted, in runs 745+746, 1-4 combination. Roughly 1000 more events are seen in x than y because of the different size diagonal cut that was used.

III. Background Extrapolation to Signal Region

Given the correct normalizations from all the data outside the xx and yy diagonal regions for the background function BK(x_l,y_l,x_r, y_r), consisting of 4 2-summed Gaussians multiplied together, one for each dimension (see above), we can now simply integrate the background function over the 4-dimensional space wherein lies the xx and yy diagonal signal region in order to find the background under the signal. We do this vdas cell by vdas cell over all 4 variables with the condition that

$$|x_l - x_r| < 3\sigma_x, \quad |y_l - y_r| < 3\sigma_y,$$

where x and y are in "effective" space (after accounting for different effective lengths on each side), and σ_x and σ_y are the left side minus right side cluster x and y difference distribution sigmas (obtained from high-t elastic events). These sigmas were about 10-15 vdas cells in magnitude, or about 0.5 - 0.8 mm. But as we do the sum over 4 dimensions, we also find the t value corresponding to each cell (i,j) left side versus cell (k,l) right side, where

$$t = 0.9^2 * [(x_i/L_{xleft} + x_k/L_{xright})/2)^2 + (y_j/L_{yleft} + y_l/L_{yright})/2)^2],$$

and L_{xleft} is the x effective length left side, etc. The background is then summed into separate t-bins that are used for the elastic data analysis. If all else were equal, we could thus do the background subtraction at this point and continue with the physics analysis of the extracted elastic signal.

However, not all else is equal. As it turns out, there is tremendous variation of cluster efficiency over x and y in our detectors. In fact, table

3 shows the expected efficiency of each x and y strip in detectors 1 (l) and 4 (r) for seeing background. The summed cluster efficiency of each strip is weighted by the background distribution within the strip. Thus, the summed cluster efficiency of each strip is equal to

$$\sum_{ijkl}^{N}(BK_{ijkl})/\sum_{ijkl}^{N}(BK_{ijkl}/\epsilon_{ij}\epsilon_{kl}),$$

where there are N cell combinations (i,j) left side and (k,l) right side in the strip region, ϵ_{ij} is the cluster efficiency of cell (i,j), etc., and BK$_{ijkl}$ is the background function contribution weight of cell pair (i,j) and (k,l). For any strip, 3 dimensions are entirely summed over (e.g. jkl) while the fourth dimension (e.g. i) is summed over the strip region only. The individual cluster efficiencies themselves are obtained from the portion of summed grey count distributions above vdas threshold at each x and y in each detector, and shall not be discussed here.

Now we wouldn't have to worry about this effect if we could assume the efficiencies were the same within the signal region as outside (where the background normalizations are done) of each x and y strip, but they are not:

vdas strip clus effs:	outside signal region	inside signal region	ratio
$50< y_l <100$	49.7 %	78.5 %	1.58
$100< y_l <150$	50.1	72.9	1.45
$150< y_l <200$	60.4	51.9	0.86
$200< y_l <220$	37.9	28.7	0.76
$120< x_l <170$	48.8	50.3	1.03
$170< x_l <220$	44.8	46.0	1.03
$220< x_l <270$	34.6	36.0	1.04
$270< x_l <320$	46.5	43.1	0.93
$320< x_l <370$	47.8	52.6	1.10
$50< y_r <100$	53.3	78.4	1.47
$100< y_r <150$	43.3	71.6	1.65
$150< y_r <180$	39.0	35.8	0.92
$180< y_r <200$	49.1	33.5	0.68
$110< x_r <160$	50.2	64.8	1.29
$160< x_r <210$	39.8	37.5	0.94
$210< x_r <260$	45.5	38.0	0.83
$260< x_r <310$	36.2	41.7	1.15

310< x_r <360	28.2	36.2	1.28.

Table 3. Summed cluster efficiencies of each x and y strip in detectors 1 (l) and 4 (r), outside and inside the signal region. The xx and yy signal regions are bounded by the diagonal regions ($\pm 3\sigma_x$ and $\pm 3\sigma_y$, respectively).

There is tremendous variation in the efficiencies of course, and more importantly, the ratios of in to out efficiencies are not equal. Hence, the normalizations of the background function as it stands would necessarily predict the wrong amount of background INSIDE the signal region, and in some cases by large amounts. To correct for this inequality, the strip ratios in column 4 of Table 3 were multiplied as "boost" factors to the background normalizations, where the strip ratios used were those corresponding to the x and y strips within which lied the vdas cell being summed. Each vdas cell being summed thus had 2 "boost" factors multiplied to its contribution from the background function, one for the xx diagonal and one for the yy diagonal. In both cases, the "boost" factor was the average of the left side strip ratio with the right side strip ratio so that boost(x) = (left x strip in/out ratio + right x strip in/out ratio)/2, and the same for y.

Finally, we have that the background function contribution of any vdas cell (i,j) left side with vdas cell (k,l) right side is:

$$\begin{aligned} BK_{ijkl} = boost(x_{ik}) * boost(y_{jl}) * N_{x(j)} * N_{x(l)} * N_{y(i)} * N_{y(k)} * \\ (G_{bg}(i) + G_{halo}(i))_{x(j)} * (G_{bg}(k) + G_{halo}(k))_{x(l)} * \\ (G_{bg}(j) + G_{halo}(j))_{y(i)} * (G_{bg}(l) + G_{halo}(l))_{y(k)}, \end{aligned}$$

where x_{ik} or y_{jl} refer to the x or y left and right side strips which cell (i,j) and cell (k,l) are within, $N_{x(j)}$ is the normalization of the x background distribution corresponding to the y strip of cell (i,j), etc., and $G_{bg}(i)_{x(j)}$ is the Gaussian x background distribution function due to beam gas at x point i, corresponding to the y strip within which lies cell (i,j), etc.

By summing over all 4 dimensions, one gets the total number of background triggers within the elastic detector combination after analysis cuts requiring clusters on each side, but BEFORE the diagonal cuts are made. Then by simply summing over all cells (i,j) left side with cells (k,l) right side that are WITHIN the xx and yy diagonal cuts, one can find the total background contribution to the signal area, and the elastic data can be extracted. Further, one can divide the background contribution into t-bin contributions (described above) for extracting the true number of elastic events, t-bin by t-bin.

IV. Results and Conclusions

As a check on this background procedure, we tried Hans Jostlein's procedure; it requires offseting the data on one side only by 1 event buffer and calculating the number of these offset-paired events found versus t in the same way the analysis is done, but now ALL these events are assumed uncorrelated background. The xx and yy diagonal analysis cuts are applied and the data are binned inside the same t-bins used in the analysis. Hence, the number of events found per t-bin should compare quite well with what the background function predicts at each t-bin, for the same data set used to get both; however, in the former case one needs to subtract from the results the percentage of real elastics in the data sample that have now been turned into uncorrelated background because of the offsetting of event records. Table 4 shows the results and comparisons for runs 745+746 (1-4 combination) where the difference between the two is included versus t. A 10 % correction to the Hans method results is applied due to the elastic "contamination" in the data file used.

	$\pm$3-sigma cut			$\pm$2-sigma cut		
t GeV/c^2	BK found w/ Hans method	BK predicted from BK ftn	diff (3σ)	BK found w/ Hans method	BK predicted from BK ftn	diff (2σ)
.0010	1986.	1956.	30.	1094.	1105.	-11.
.0017	923.	816.	108.	448.	427.	21.
.0024	466.	425.	42.	209.	220.	-11.
.0031	276.	260.	16.	134.	129.	6.
.0038	175.	173.	2.	71.	84.	-13.
.0045	124.	121.	3.	60.	59.	1.
.0052	80.	88.	-8.	33.	44.	-10.
.0059	76.	66.	10.	41.	33.	8.
.0066	63.	50.	13.	31.	25.	6.
.0073	45.	38.	7.	13.	19.	-7.
.0080	31.	29.	2.	13.	15.	-2.
.0087	27.	22.	5.	12.	11.	0.
.0094	24.	17.	8.	13.	9.	4.
.0101	14.	12.	1.	8.	7.	2.

.0108	12.	9.	3.	7.	5.	2.
.0115	7.	6.	1.	4.	4.	1.
.0122	6.	4.	2.	3.	2.	0.
.0129	4.	3.	1.	2.	2.	0.
.0136	1.	2.	-1.	1.	1.	0.
TOTALS	4342.	4097.	245.	2198.	2201.	-3.

Table 4. Comparisons of background predicted by the background function versus Hans method in runs 745+746 (1-4 combination).

For the comparisons in Table 4 to be valid, one must assume uncorrelated background between the two sides and between x and y itself on each side, which is not an entirely good assumption. Also, while cluster efficiency corrections need not be applied to Hans method, the cluster efficiency ratio corrections (see above) MUST be applied to the background function in order to fairly extrapolate the normalization region underneath the elastic signal region; since these ratios have a 10 % or more error in runs 745+746 (1-4 combination), the comparisons are actually quite good. In fact, only a few (admittedly low-t) bins have much disagreement at all.

Table 5 lists the latest counts of signal versus background events for low-t runs 745+746 (no main ring), the 1-4 combination. The analysis with a $\pm$3-sigma cut along the xx and yy diagonals is compared against that with a $\pm$2-sigma cut. These backgrounds are what the background function predicts to be INSIDE both diagonals. The total data found inside the diagonals at each t point is the sum of the elastics and bckgrd columns. Again, the 1-4 elastic combination of tdcs (heel or toe) is required with no ring counters in time and nothing in time in detectors 2 and 3. Also, the heel adc above threshold is required in both detectors 1 and 4.

	$\pm$3-sigma cut		$\pm$2-sigma cut	
t	Elastics	Bckgrd	Elastics	Bckgrd
.0010	391	1956	227	1105
.0017	315	816	169	427
.0024	173	425	125	220
.0031	138	260	102	129
.0038	141	173	99	84
.0045	81	121	79	59
.0052	83	88	88	44

.0059	106	66	100	33
.0066	115	50	118	25
.0073	111	38	108	19
.0080	95	29	93	15
.0087	120	22	121	11
.0094	100	17	98	9
.0101	90	12	91	7
.0108	65	9	64	5
.0115	70	6	66	4
.0122	50	4	48	2
.0129	47	3	46	2
.0136	31	2	30	1
TOTALS	2322	4097	1872	2201.

Table 5. Comparison of signal to background at low-t for runs 745+746, 1-4 combination.

One can see that the ratio of signals of the $\pm$2-sigma cut case to $\pm$3-sigma cut case is much less than 97 %, leading to the conclusion that there is still some background contamination in the $\pm$3-sigma cut case. Therefore, further study on possible contamination needs to be done. But what may be problematic is that these numbers have a strong dependence on the cluster efficiency maps of each detector; and since these efficiencies were obtained with low statistics and with very crude bin sizes, we may be at the point of no return until more statistics on cluster efficiency data becomes available.

Having mentioned this however, one notices that the signal to background ratio of 1/1 is reached just above t=0.005 $(\mathrm{GeV}/c)^2$, and that the total summed signal to background ratio over the active t-range of the 1-4 combination in runs 745+746 is slightly better than 1 to 2 after all analysis cuts are made. We will clearly want to do better than this in future runs.