Weekly Update (Christmas Edition)

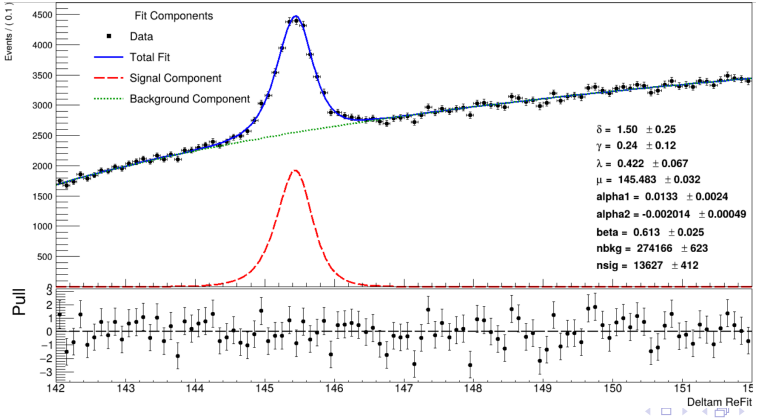
Edward Wardell

UoE Christmas Period

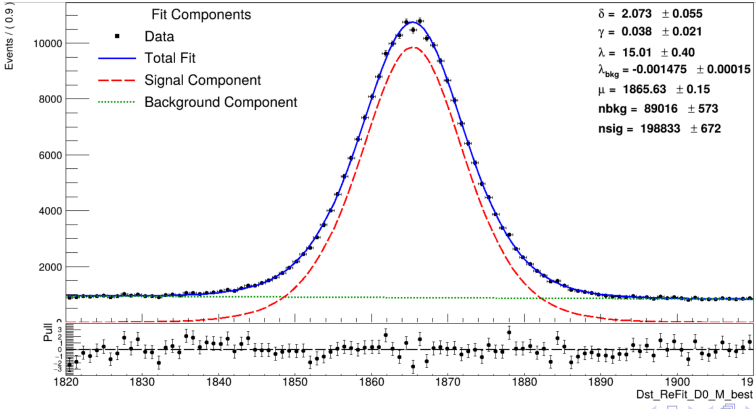
Tasks Undertaken

- Obtained mass fits.
- Methods to apply RS clusters to WS data.

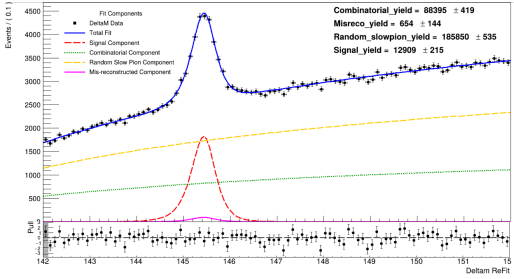
WS Mass Fit of deltam_ReFit: (Signal = Johnson), (Background = Polynomail)

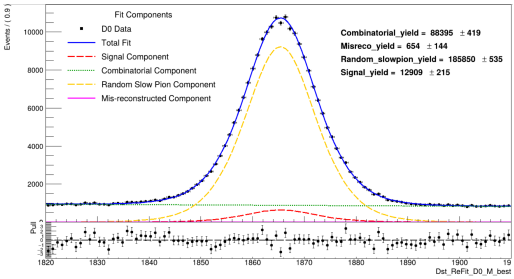


WS Mass Fit of Dst_ReFit_D0_M_best: (Signal = Johnson), (Background = Negative Exponential)

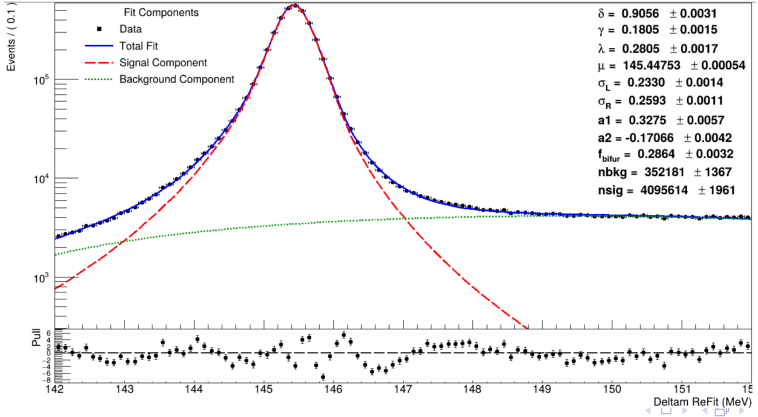


WS 2D Mass Fit - Finding Yields

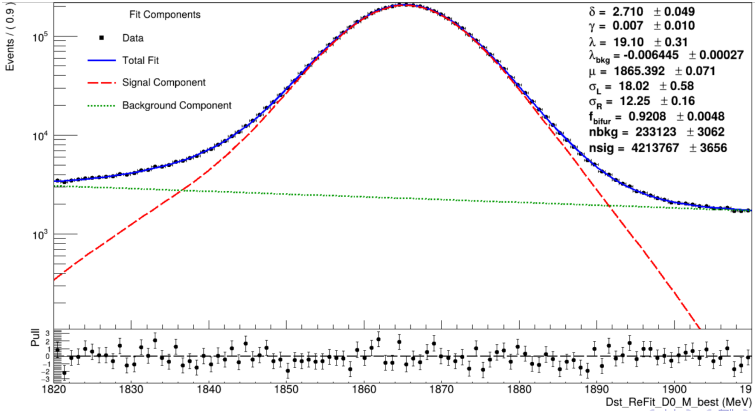




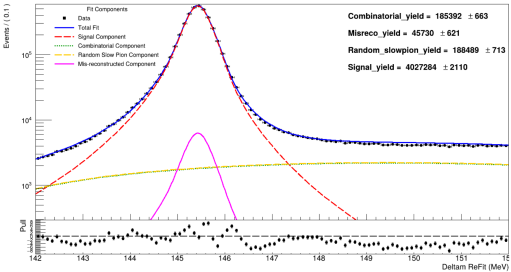
RS Mass Fit of deltam_ReFit: (Signal = Johnson + Bifurcated Gauss), (Background = Chebychev)

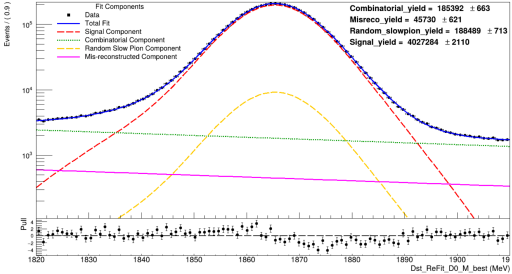


RS Mass Fit of Dst_ReFit_D0_M_best: (Signal = Double Gauss + Bifurcated Gauss), (Background = Negative Exponential)

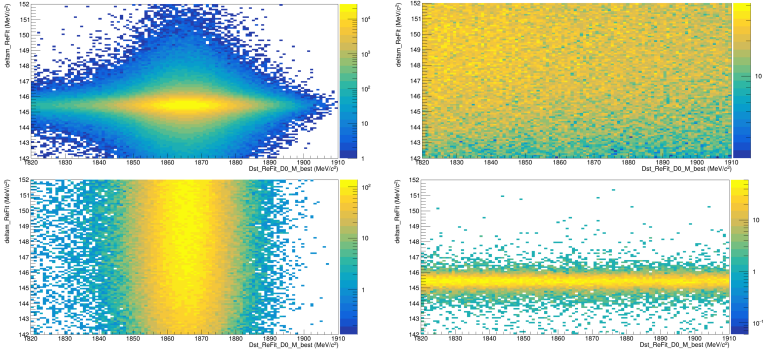


RS 2D Mass Fit - Finding Yields



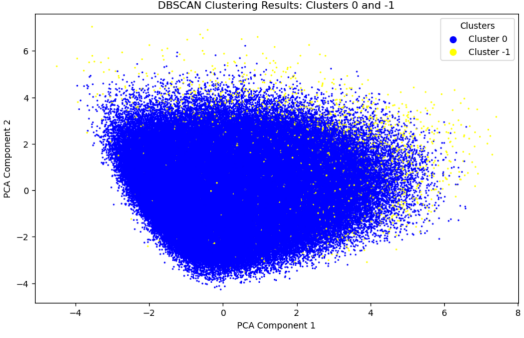


Signal and Background PDFs - 2D plots



Unsupervised: 10D DBSCAN, Applying RS to WS with "Centroid" method.

RS Cluster with DBSCAN

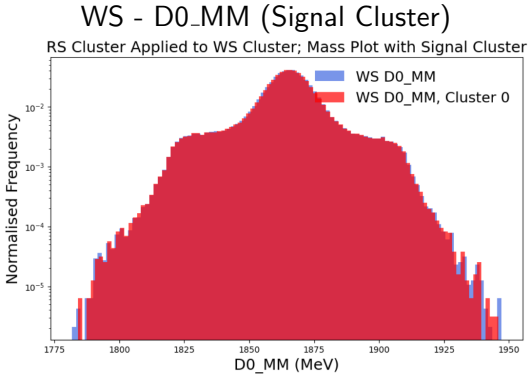


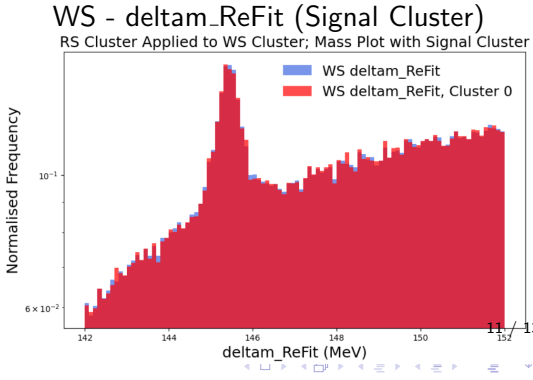
WS Clusters with "Centroid" distance

cluster_labels	
0	196266
-1	91532

cluster_labels

Unsupervised: 10D DBSCAN, Applying RS to WS with "Centroid" method.





K-Means Clustering & Optimum number of clusters in 10D.

- How K-Means Works
- Input: D: The dataset containing the points $[x_1, x_2, ..., x_N]$, where N is the total number of data points. K: The desired number of clusters.
- Choose K initial centroids $[c_1, c_2, ..., c_K]$ randomly from the dataset. These centroids represent the initial "centre" of each cluster.
- Now find the nearest centroid id.
- Then assign points to cluster.
- Re-calculate centre of cluster until convergence and cluster assignments do not change.

K-Means Proprties

- Minimises aggregate intra-cluster distance: $V = \sum_k \sum_{x_n \to c_k} ||x_n - c_k||^2$.
- In Euclidean distance, then V is proportional to variance.
- Find an "appropriate" K: optimise for V.