CDF/ANAL/BOTTOM/CDFR/10334 Version 4.0
May 10, 2011

Observation of $B_s \to J/\psi f_0(980)$ decays

Michal Kreps[1],

Physics Department, University of Warwick

Abstract

We present observation of the decay $B_s \to J/\psi f_0(980)$ with $f_0(980) \to \pi^+\pi^-$. Using 3.8 fb^{-1} of dimuon data we select sample of events with clear B_s signal, which is consistent with studied decay. Using this signal, we measure ratio of branching fractions relative to $B_s \to J/\psi\phi$ to be

$$R = \frac{\mathcal{B}(B_s \to J/\psi f_0(980))}{\mathcal{B}(B_s \to J/\psi\phi)} \frac{\mathcal{B}(f_0(980) \to \pi^+\pi^-)}{\mathcal{B}(\phi \to K^+K^-)} = 0.257 \pm 0.020 \pm 0.014.$$

[1]M.Kreps@warwick.ac.uk

Contents

1 Introduction

Since B_s mixing observation in 2006 [1] particle physics community is excited by searches for CP violation in $B_s \to J/\psi\phi$ decays. First results which used flavour tagging to separate B_s and $\overline{B}_s$ from late 2007 and early 2008 showed about 1.5 σ deviations from the standard model [2, 3]. In both initial measurements as well as in the subsequent CDF update [4] possible contribution from the so-called s-wave contribution was neglected. The s-wave component itself can be non-resonant $B_s \to J/\psi K^+K^-$ decay or decay $B_s \to J/\psi f_0(980)$ with $f_0(980) \to K^+K^-$. This neglection caused significant discussion which mainly boiled down to the question whether the departure of Tevatron data from the standard model are due to bias caused by neglecting s-wave contribution or the effect is genuin. A summary of this discussion can be find in Refs. [5, 6]. It is also suggested that sufficient signal for decay $B_s \to J/\psi f_0(980)$ with $f_0(980) \to \pi^+\pi^-$ can be used to measure CP violationg phase $\beta_s^{J/\psi\phi}$ without need of angular analysis [7]. All this generated interest in the decay $B_s \to J/\psi f_0(980)$. Given the large dataset we have at the CDF it is interesting to search for this decay.

In this note we report observation of the $B_s \to J/\psi f_0(980)$ decay with $f_0(980) \to \pi^+\pi^-$ and measurement of its branching fraction using 3.8 fb^{-1} of CDF data. The note is organized as follows:

2 Data and simulation samples

In this analysis we use data collected between February 2002 and August 2008, in CDF callibration periods this corresponds to periods 0 to 20. The integrated luminosity of the sample is 3.8 fb^{-1}. We use events collected by the J/ψ trigger which ends in the xpmm dataset. For analysis we use BottomMods package with custom tcl file to reconstruct decay $B_s \to J/\psi f_0(980)$ as well as normalization decay $B_s \to J/\psi\phi$.

The analysis uses several simulated samples, all of them being simulated using BGen or HeavyQuarkGen. In each sample we produce single b-hadron which is then decayed by EvtGen to wanted decay channels. Samples typically simulate first 780 pb^{-1} of our dataset. First samples simulates B_s mesons, which is forced to decay to $J/\psi f_0(980)$ final state with $J/\psi \to \mu^+\mu^-$ and $f_0(980) \to \pi^+\pi^-$. Second sample is analoug of the first one, but B_s decays to $J/\psi\phi$ with $\phi \to K^+K^-$. Two other samples provide attempt to generic type of simulation. Each of them generates B^0, B^+, B_s and Λ_b with inclusive decay to final states containing J/ψ. First of these was generated for the $Y(4140)$ analysis [8], while second one by the author of the analysis. Those samples are used to understand possible physics backgrounds. Last sample is again pure B_s sample, but this time it decays to three decay channels, $B_s \to J/\psi\phi$ and $B_s \to J/\psi f_0(980)$ with $f_0(980)$ decaying to both pair of charged pions and pairs of charged kaons. This last sample is independent of first one used to derive selection and is used to determine reconstruction efficiencies needed. All samples are reconstructed using same tcl as used for data.

2.1 Simulation weighting

The simulated samples we use are generated using phase space decay model. To evaluate efficiencies and resolutions we weigh the samples for best knowledge of physics of given decay.

We choose this aproach as it allows us easy resampling with different physics parameters, which is useful to evaluate systematic uncertainties.

To weigh $B_s \to J/\psi\phi$ decay we use following probability density function

$$\begin{aligned}\frac{\mathrm{d}^4P(\vec{\rho},t)}{\mathrm{d}\vec{\rho}\mathrm{d}t} &= |A_0|^2 f_1(\vec{\rho})\tau_+ + |A_{||}|^2 f_2(\vec{\rho})\tau_+ + |A_\perp|^2 f_3(\vec{\rho})\tau_- \\ &+ |A_0||A_{||}| f_5(\vec{\rho})\cos(\delta_{||})\tau_+.\end{aligned}$$

Here we assume that CP is conserved which removes other two interference terms. The functions $\tau_\pm$ and $f_i(\vec{\rho})$ are defined as

$$\begin{aligned}\tau_+ &= e^{-\Gamma_L t}, \\ \tau_- &= e^{-\Gamma_H t}, \\ f_1(\vec{\rho}) &= \frac{9}{32\pi} 2\cos^2\psi\left(1-\sin^2\theta\cos^2\phi\right), \\ f_2(\vec{\rho}) &= \frac{9}{32\pi}\sin^2\psi\left(1-\sin^2\theta\sin^2\phi\right), \\ f_3(\vec{\rho}) &= \frac{9}{32\pi}\sin^2\psi\sin^2\theta, \\ f_4(\vec{\rho}) &= -\frac{9}{32\pi}\sin^2\psi\sin 2\theta\sin\phi, \\ f_5(\vec{\rho}) &= \frac{9}{32\pi}\frac{1}{\sqrt{2}}\sin 2\psi\sin^2\theta\sin 2\phi, \\ f_6(\vec{\rho}) &= \frac{9}{32\pi}\frac{1}{\sqrt{2}}\sin 2\psi\sin 2\theta\cos\phi,\end{aligned}$$

with $\vec{\rho} = (\theta, \psi, \phi)$ are angles defined in transversity basis. The parameters including their uncertainties are taken from CDF β_s measurement [9] except $\phi_{||}$ which is not determined in that analysis and which value we take from PDG [10]. Their values are as follows

$$\begin{aligned}\tau &= 458.6 \pm 8.4\ \mu\mathrm{m}, \\ \Delta\Gamma &= 0.075 \pm 0.036\ \mathrm{ps}^{-1}, \\ |A_0|^2 &= 0.524 \pm 0.020, \\ |A_{||}|^2 &= 0.231 \pm 0.021, \\ \phi_{||} &= -2.86 \pm 0.11, \\ \phi_\perp &= 2.95 \pm 0.64.\end{aligned}$$

We need also $B^0 \to J/\psi K^{*0}$ weighting as $B^0 \to J/\psi K^{*0}$ is physics background for $B_s \to J/\psi\phi$ which we need to take into account. To weigh $B^0 \to J/\psi K^{*0}$ we use following probability density function

$$\begin{aligned}\frac{\mathrm{d}^4P(\vec{\rho},t)}{\mathrm{d}\vec{\rho}\mathrm{d}t} &= |A_0|^2 f_1(\vec{\rho}) + |A_{||}|^2 f_2(\vec{\rho}) + |A_\perp|^2 f_3(\vec{\rho}) \\ &+ |A_0||A_{||}| f_5(\vec{\rho})\cos(\delta_{||}) \\ &\pm |A_{||}||A_\perp| f_4(\vec{\rho})\sin(\delta_\perp - \delta_{||}) \\ &\pm |A_0||A_\perp| f_6(\vec{\rho})\sin\delta_\perp\end{aligned}$$

where $\pm$ sign corresponds to the charge of kaon from K^* decay. The parameters are taken from PDG [10] to be

$$\begin{aligned} |A_0|^2 &= 0.570 \pm 0.008, \\ |A_\perp|^2 &= 0.219 \pm 0.010, \\ \phi_{||} &= -2.86 \pm 0.11, \\ \phi_\perp &= 3.01 \pm 0.14. \end{aligned}$$

3 Candidate reconstruction

The reconstruction of the candidates starts with track reffiting in which if available L00 hits are attached to the tracks and appropriate multiple scattering and energy loss is taken into account. As multiple scattering and energy loss depends on the mass of the particle, each track is fitted with muon, pion and kaon mass hypothesis. After track reffiting, pairs of muons of opposite charge form J/ψ candidates. Candidates with invariant mass between 2.85 and 3.25 GeV/c^2 and fit χ^2, quality of a kinematical fit with constraint that both tracks come from same vertex, being less than 30 are retained for further steps.

In events with J/ψ candidate, we continue by reconstructing $f_0(980)$ and ϕ by combining tracks of opposite charge. The $f_0(980)$ is reconstructed from the tracks using pion mass hypothesis while ϕ tracks use kaon mass hypothesis. Candidates are subject to kinematical fit with vertex constraint. We require that χ^2 of the kinematical fit is less than 30 together with requirement on invariant mass of the candidate to be between 0.85 and 1.2 GeV/c^2 for $f_0(980)$ and 0.98 and 1.2 GeV/c^2 for ϕ.

Last step combines J/ψ candidates with $f_0(980)$ and ϕ candidates to form B_s candidates. Each candidate is subjet to four track kinematic fit with constraint that all tracks originate from same vertex and dimuon invariant mass being equal to world average J/ψ mass. Candidates with χ^2 less than 50 and invariant mass between 4.8 and 6.2 GeV/c^2 are retained for the analysis.

4 Candidate selection

The selection of the candidates starts with impossing few requirements, which aims in removing large fraction of background while keeping most of the signal. We require that each track has transverse momentum of at least 400 MeV/c, at least 10 axial and at least 10 stereo hits inside COT and at least 3 axial hits in silicon tracker. In addition we require transverse momentum of B_s candidate to be larger than 4 GeV/c, significance of transverse displacement of the secondary vertex from primary vertex to be larger than 3 and invariant mass of the dimuon to be between 3.0 and 3.2 GeV/c^2. We also tighten B_s candidate invariant mass range to 5.1 to 5.6 GeV/c^2. For normalization mode $B_s \to J/\psi\phi$ we also restrict kaon pair invariant mass to interval between 1.009 and 1.0029 MeV/c^2. For signal mode $B_s \to J/\psi f_0(980)$ we keep wide pion pair invariant mass window in order to allow possibility of cross-checking that it is consistent with expectation and to avoid large systematic uncertainties for selection efficiency as mass shape of $f_0(980)$ is not very precisely known. Finally we apply official good run list provided by DQM group.

Input	rank	id	add signi	only this	loss	global corr.
$Pt(h)$	1	5	533.24	533.24	236.35	71.7%
$\chi^2(B_s)$	2	2	275.42	495.18	81.29	90.8%
$ct(B_s)$	3	8	127.81	280.47	127.96	26.5%
$Prob(B_s)$	4	4	46.96	463.35	46.04	91.4%
$\cos\theta_\pi$	5	10	28.75	92.25	35.45	25.9%
$Pt(B_s)$	6	3	33.16	322.13	32.46	57.8%
$Prob(h)$	7	6	17.47	48.48	17.49	33.1%
$\cos\theta_\mu$	8	9	3.77	208.34	3.78	43.0%
$Prob(J/\psi)$	9	7	1.35	106.17	1.35	40.7%

Table 1: List of inputs to neural network with rank providing importance, id gives column/row in Fig. 1, add signi amount of information this input adds to the overal performance. Only this means amount of information which single input carries, loss corresponds to amount of information which is lost when input is removed and global corr. provides information on the global correlation of given input to all other inputs.

The final candidate selection is performed using neural network. The neural network is build using NeuroBayes package [11, 12] and designed to be universal for both signal and normalization decays. This means that only quantities which can be defined for both decays are used as inputs into the neural network. In the following by the name hadron we mean $f_0(980)$ or ϕ mesons in the B_s candidate. The inputs into the neural network are:

- χ^2 of the R-ϕ projection of the kinematical fit of B_s candidate (2). This input is second most important.
- Transverse momentum of the B_s candidate (3).
- Probability derived from the full χ^2 of the kinematical fit of B_s (4).
- Transverse momentum of hadron (5). This is the most important input.
- Probability derived from the full χ^2 of the kinematical fit of hadron (6).
- Probability derived from the full χ^2 of the kinematical fit of J/ψ (7).
- Proper decay time multiplied by c of B_s candidate. It is calculated as product of transverse displacement and reconstructed invariant mass divided by transverse momentum (8). It is third most important input to the neural network.
- Cosine of helicity angle of positive muon (9) and positive pion/kaon (10). Cosine of muon is calculated as a normalized dot product of muon three momentum in J/ψ rest frame and J/ψ momentum in B_s rest frame. Calculation for pion/kaon is analogous.

The numbers in parenthesis correspond to the number in some of the figures we show to demonstrate main features of the neural network. The neural network was trained using about 230k simulated signal events and as background about 450k data events taken from B_s invariant mass region 5.45 to 5.55 GeV/c^2.

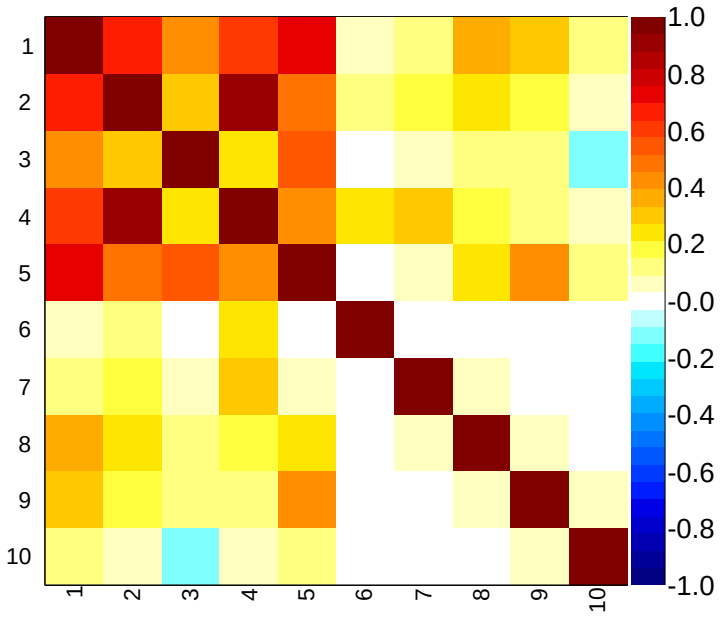

Figure 1: Correlation matrix of the inputs of neural network derived from the training sample. Column/row number 1 represents truth, e.g. whether candidate is signal or background. Other columns/rows show inputs, with coding corresponding to the numbers in parethesis in inputs list.

To get a visual impression of the correlation between different variables and the fact whether candidate is signal or background, in Fig. 1 we show correlation matrix derived on the training sample. In Fig. 2 we show distribution of the neural network output separately for signal and background. In Fig. 2 we also show purity, number of signal events over number of all events, in bins of neural network output. If neural network is well trained, this should show linear dependance with unit slope, which is case here. To decide where to place requirement on the neural network output, we compute $F = N_S/(2.5 + \sqrt{N_b})$ for different requirements and choose one which maximizes this quantity. The N_S is number of candidates selected in simulation and N_b is number of background events extrapolated from the B_s mass sidebands. At this stage we also remove few low statistics L2 trigger bits to achieve better agreement between data and simulation. List of accepted L2 triggers is in table 2. In Fig. 3 we show dependance of F on the requirement on neural network output. Please note that scale on the y-axis is arbitrary and does not provide any information on significance we expect. From dependance itself we select candidates which have neural network output greater than 0.96. Together with it, we also require that two muons match dimuon trigger primitives as determined by the `MuonTrigMatch` class. Same selection is applied also to normalization $B_s \to J/\psi\phi$ decay. In Fig. 4 we show the invariant mass distributions of selected candidates both for signal and normalization channels. Clear B_s signal is observed in $J/\psi\pi^+\pi^-$ mass spectrum.

We select 9914 $B_s \to J/\psi f_0(980)$ candidates with 528 candidates coming from events with

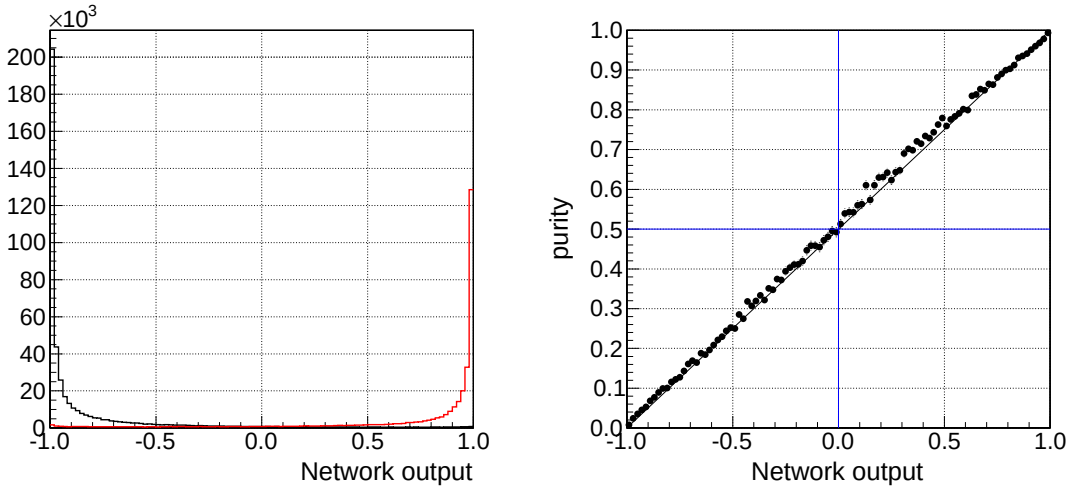

Figure 2: Distribution of the neural network output for signal (red) and background (black) using training sample (left). Purity in bins of neural network output on training sample (right).

class 1
L2_AUTO_L1_CMU1.5_PT1.5_&_CMX1.5_PT2_CSX
L2_PS100_L1_CMU1.5_PT1.5_&_CMX1.5_PT2_CSX
L2_CMU1.5_PT1.5_&_CMX1.5_PT2_DPHI120_OPPQ
L2_CMU1.5_PT1.5_&_CMX1.5_PT2_DPHI120_OPPQ_DPS
class 2
L2_AUTO_L1_TWO_CMU1.5_PT1.5
L2_PS100_L1_TWO_CMU1.5_PT1.5
L2_TWO_CMU1.5_PT1.5_DPHI120_OPPQ_DPS
class 3
L2_CMUP1.5_PT3_&_CMU1.5_PT1.5
L2_CMUP1.5_PT3_&_CMU1.5_PT1.5_DPS
L2_CMU1.5_PT1.7_&_CMU1.5_PT3_1.7MT7_DPS
L2_CMU1.5_PT1.7_&_CMU1.5_PT3_1.7MT7_LUMI_185
L2_CMU1.5_PT1.7_&_CMX1.5_PT3_1.7MT7_DPS

Table 2: L2 triggers selected for the analysis. We split them into three classes based on the actual requirements of the trigger. When even falls into more than one class, it is kept in class with higher number.

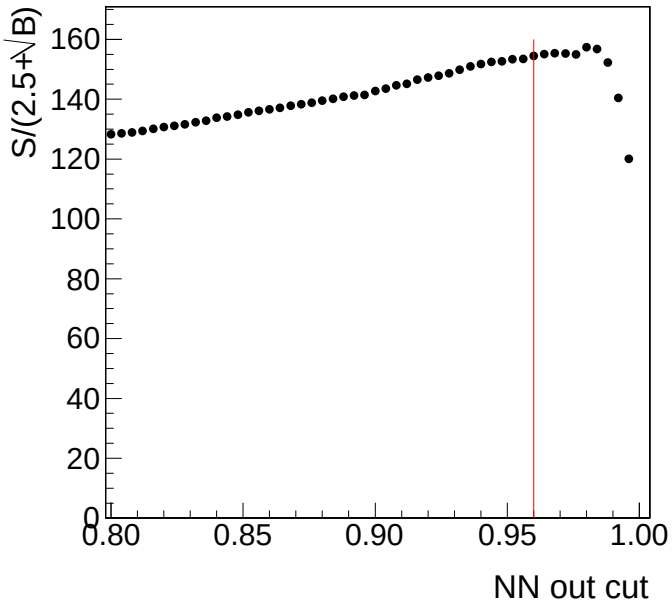

Figure 3: Dependance of the figure of merit, $F = N_S/(2.5 + \sqrt{N_b})$, on neural network output requirement. Each point corresponds to selection of candidate wich neural network output larger than given value. Red vertical line shows final requirement.

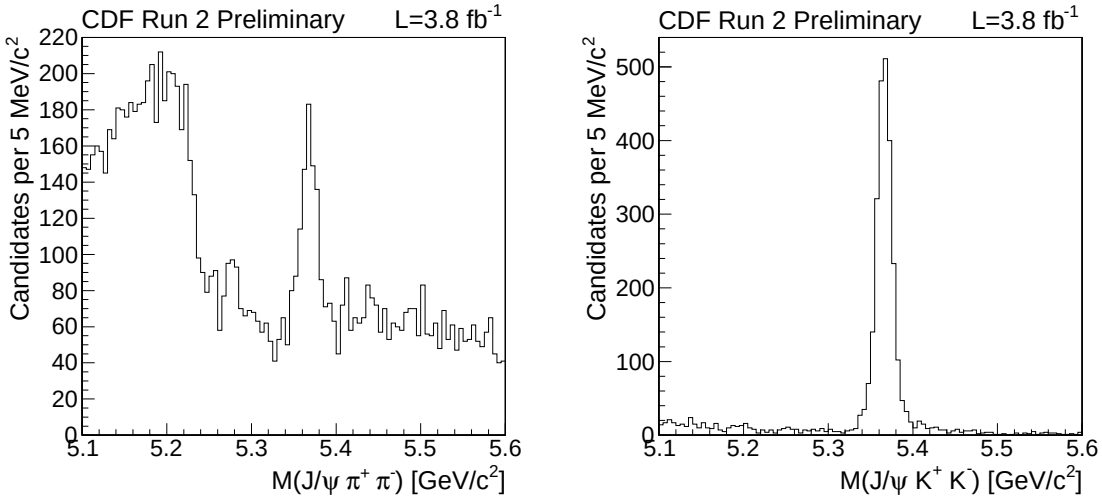

Figure 4: Invariant mass distribution of selected $B_s \to J/\psi f_0(980)$ candidates (left) and of $B_s \to J/\psi\phi$ candidates (right).

more than one candidate selected. Out of the events with multiple candidates, 229 of them have two candidates selected and 20 more than two candidates. In normalization sample of $B_s \to J/\psi\phi$ decays we select 3059 candidates. There are 7 events in sample yielding two candidates per event. Here we can also evaluate effect of restricted trigger selection, which removes about 100 events from $B_s \to J/\psi\phi$ sample. Given that normalization mode has larger signal statistics than our signal mode and that we don't observe multiple candidates in events within normalization sample we conclude that it is unlikely that events with multiple candidates would give more than one signal event. This argument together with fact that we generate only single b-hadrons to evaluate efficiency and thus we cannot check determine efficiency of any selection which would try to remove multiple candidates, we keep all of them in the sample. We also explicitly check that none of the events with multiple candidates contains single candidate more than once.

5 Physics Backgrounds

Before going into the detailed study of significance and branching fraction for observed signal it is important to check that it does not arise from some other, already known decay. For this we use generic type sample where b-hadrons are forced to decay to any possible decay mode containing J/ψ. This is good approximation given the cleanest of the J/ψ sample at CDF. In Fig. 5 we show invariant mass distributions of events reconstructed in $J/\psi\pi^+\pi^-$ channel in the generic J/ψ simulation. Based on it we can conclude that B^+ and Λ_b do not contribute any significant background in our data. We also see that while B^0 and B_s contribute with some events, they do not peak at the B_s mass itself. It is interesting to note that structure seen in Fig. 5 for B^0 and B_s can at least qualitatively describe the broad structure at lower masses in Fig. 4.

We also investigate in more details composition of the physics background stemming from B^0 and B_s decays. In B^0 sample we identify two outstanding decays, namely $B^0 \to J/\psi K^{*0}$ with mis-reconstructed K^{*0} and $B^0 \to J/\psi\rho^0$ with ρ^0 decaying to pair of charged pions. This second decay is properly reconstructed and is the origin of small peak at B^0 mass in Fig. 5. In Fig. 6 we show invariant mass distributions of the two identified decays together with distribution for all other decays. Same study on B_s sample identifies only single rather string decay which is nonresonant $B_s \to J/\psi K^+K^-$. Decay $B_s \to J/\psi\phi$ with $\phi \to K^+K^-$ does not contribute significantly. One potentially dangerous decay identified is $B_s \to J/\psi\phi$ with $\phi \to \mu^+\mu^-$ which is almost correctly reconstructed due to the closenest of pion and muon masses. On the other hand, smallnes of $\phi \to \mu^+\mu^-$ branching fraction renders this contribution to be negligible. Figure 7 shows invariant mass distributions of physics background stemming from B_s split to different decay channels. While there might be indication that $B_s \to J/\psi\phi$ decay could peak close to B_s mass, its size is negligible. To check this we take much larger sample of simulated $B_s \to J/\psi\phi$ decays used in β_s analysis and reconstruct it both as $B_s \to J/\psi\phi$ and $B_s \to J/\psi f_0(980)$. In this sample we correctly reconstruct 623620 events while we reconstruct 107 events as $B_s \to J/\psi f_0(980)$, which translated to relative efficiency of $1.72 \cdot 10^{-4}$. Taking into account size of the selected normalization decay events (about 3156 events) we expect maximum of about 0.5 events in our signal sample. This is definitely far below any reasonable sensitivity of current analysis and so we don't attempt to estimate uncertainty on this contribution as even factor of 2 underestimation would not make significant difference.

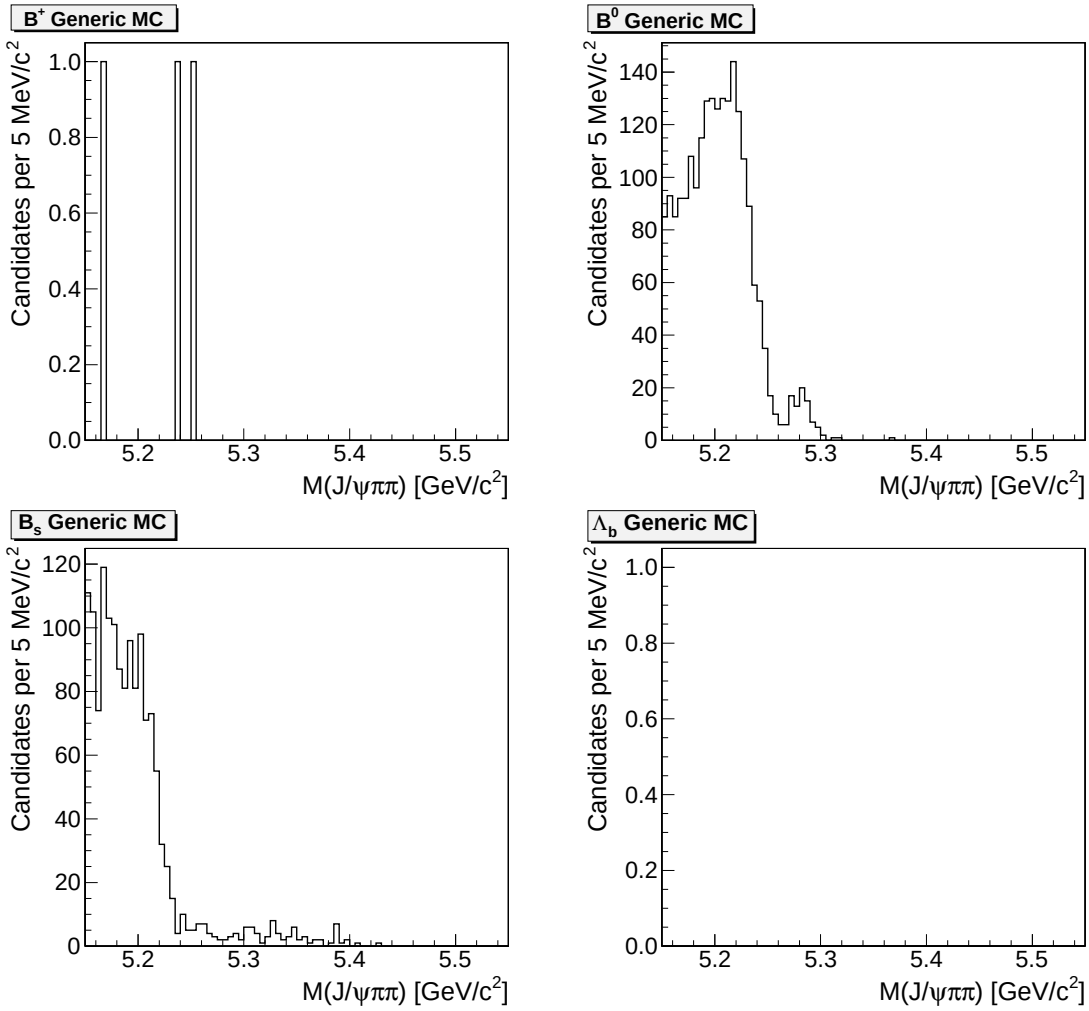

Figure 5: Contribution of physics backgrounds to signal $B_s \to J/\psi f_0(980)$ decay determined from the generic simulation sample separated to B^+ (top left), B^0 (top right), B_s (bottom left) and Λ_b (bottom right).

One decay not mentioned here is self-reflection of $B_s \to J/\psi f_0(980)$ with $f_0(980) \to K^+K^-$ which would be misreconstructed. In first order kinematics of such events is similar to $B_s \to J/\psi\phi$ decay and thus it should have also similar invariant mass distribution and relative efficiency. Taking into account also branching fractions, we would expect almost order of magnitude smaller contribution than already neglected misreconstructed $B_s \to J/\psi\phi$.

Similar study can be made also for normalization channel. Figure 8 shows $J/\psi K^+K^-$ invariant mass distributions from generic simulation sample without excluding decay $B_s \to J/\psi\phi$. Again, B^+ and Λ_b do not contribute with significant background. In case of B_s it turns out that whole reconstructed signal is due to $B_s \to J/\psi\phi$ and nonresonant $B_s \to J/\psi K^+K^-$ decays. While it would be interesting to separate those, it would require dedicated study on data, which is beyond scope of this analysis. Therefore we will treat all events as coming from $B_s \to J/\psi\phi$. From B^0 decays, only well known $B^0 \to J/\psi K^{*0}$ contributes and it peaks close to the B_s mass. This misreconstruction effect is well known since early studies of

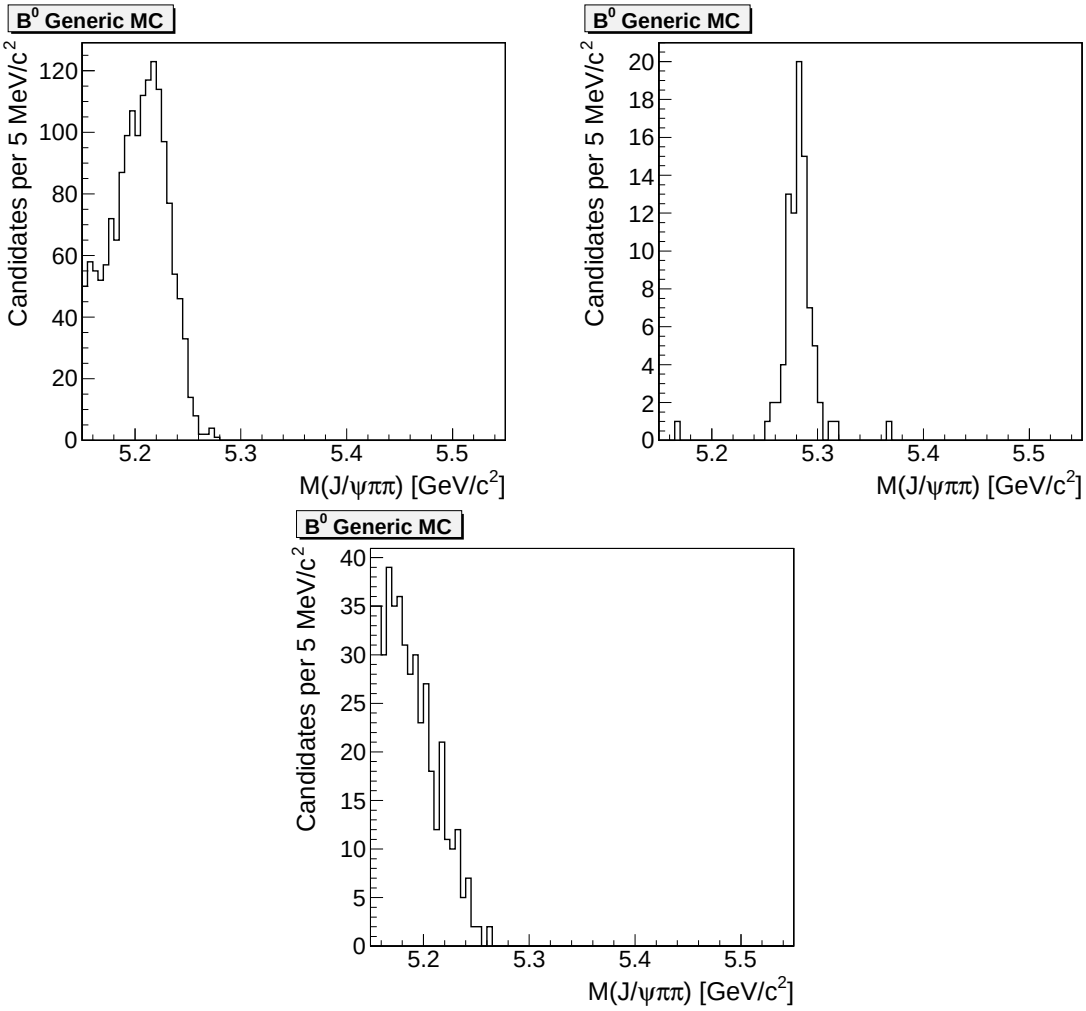

Figure 6: Split of the physics background to $B_s \to J/\psi f_0(980)$ stemming from B^0 decays. Top left shows invariant mass distribution of decays $B^0 \to J/\psi K^{*0}$, top left of decays $B^0 \to J/\psi \rho^0$ and bottom plot shows all other decays.

$B_s \to J/\psi\phi$ decays at CDF. As quick first estimate gives size of this contribution of about 3% of $B_s \to J/\psi\phi$ we will properly treat it in the fit in order to keep systematic uncertainties under control.

6 Fit description

To extract number of events we use unbinned extended maximum likelihood fit of the invariant mass. While there are some differences between signal $B_s \to J/\psi f_0(980)$ and normalization $B_s \to J/\psi\phi$ decays, the two fits have also many similarities and therefore we implement single likelihood function which is used for both decay channels. In this section we describe details of the fit including inputs which are made, fit validation. At end of the section we extract number of $B_s \to J/\psi f_0(980)$ and $B_s \to J/\psi\phi$ events used to extract branching fraction ratio.

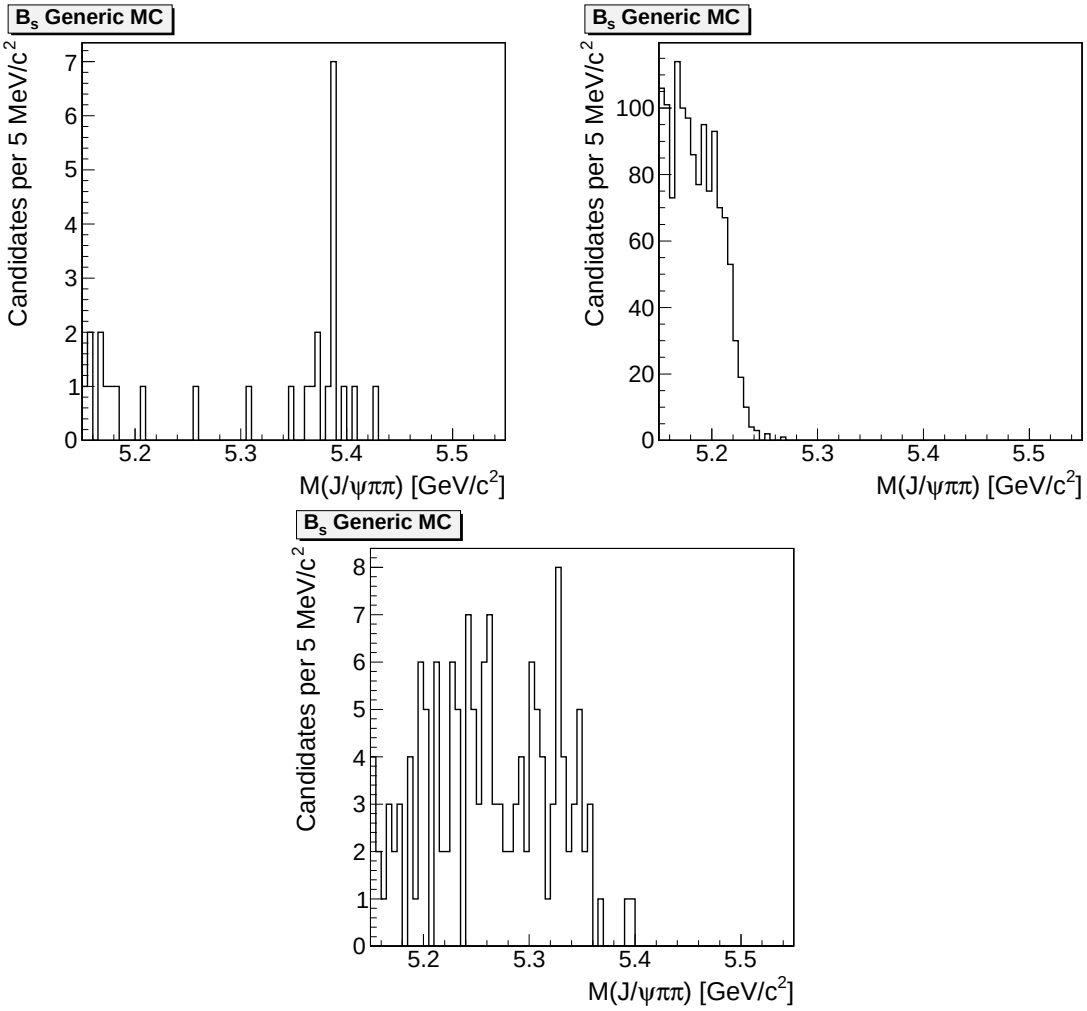

Figure 7: Split of the physics background to $B_s \to J/\psi f_0(980)$ stemming from B_s decays. Top left shows invariant mass distribution of decays $B_s \to J/\psi\phi$ with $\phi \to K^+K^-$, top right is for nonresonant decay $B_s \to J/\psi K^+K^-$ and bottom plot shows all other decays.

6.1 General likelihood

The likelihood used in the fit has form

$$\mathcal{L} = \prod_{i=1}^{N} [N_s \cdot P_s(m_i) + N_b \cdot P_b(m_i) + f_{phb} \cdot N_s \cdot P_{phb}(m_i) + N_{B^0} \cdot P_s(m_i)] \cdot e^{-(N_s+N_b+N_s \cdot f_{phb}+N_{B^0})}, \quad (1)$$

where N is the total number of candidates in the sample, N_s and N_b are number of signal and background events and m_i is the invariant mass of i-th candidate. First two terms represent signal and combinatorial background. Term with subscript *phb* corresponds to physics backgrounds and as their size is generally know with respect to signal, we bound it also in likelihood to the signal. Last term with density function $P_s(m_i)$ adds another signal

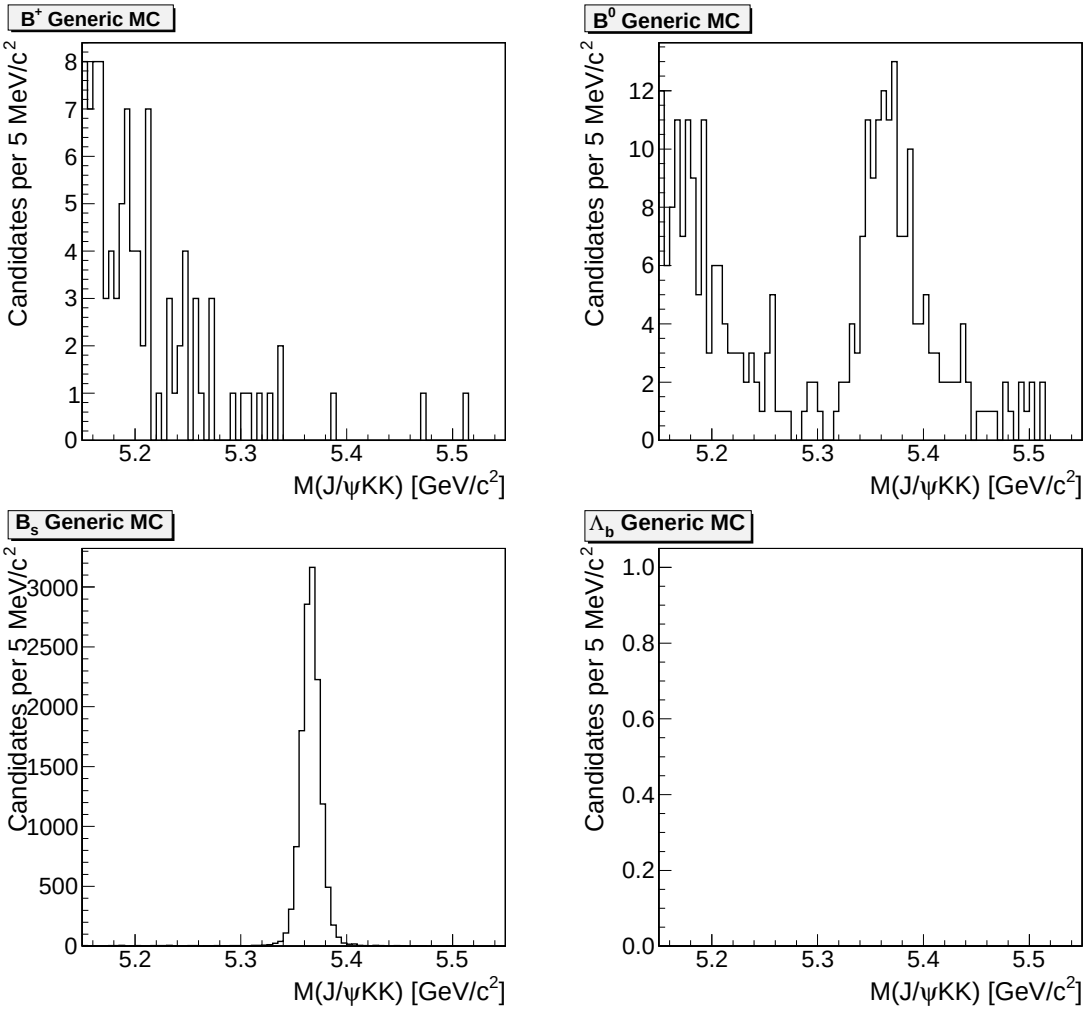

Figure 8: Contribution of physics backgrounds to normalization $B_s \to J/\psi\phi$ decay determined from the generic simulation sample separated to B^+ (top left), B^0 (top right), B_s (bottom left) and Λ_b (bottom right).

like shape to the background to cope with contribution of correctly reconstructed B^0 decays. Mass shape is assumed to be same as for $B_s \to J/\psi f_0(980)$ signal with position fixed to 5.279 GeV/c^2. Number of B^0 events N_{B^0} is left free in the fit. The signal probability density function is parametrized by tripple Gauss function with exact form being

$$P_s(m_i) = f_1 \cdot G\left(M, t\sigma_1\right) + \left(1 - f_1\right) f_2 \cdot G\left(M, t\sigma_2\right) + \left(1 - f_1\right)\left(1 - f_2\right) \cdot G\left(M, t\sigma_3\right). \qquad (2)$$

Here f_1 and f_2 denote fractions between three Gaussians, σ_i is the width of i-th Gaussian, M is the mean B_s mass and t is single scaling factor for widths of the Gaussians to incorporate possible differences in overal scale of the resolution between simulation and real data. While three Gaussians are implemented we find in $B_s \to J/\psi f_0(980)$ simulation that third Gaussian is insignificant. Therefore for both decays we use just two Gaussians to describe signal shape. The probability density function $P_b(m_i)$ for the combinatorial background is parametrized by Chebyshev polynomials of the first kind. The physics background is parametrized by the

probability density function $P_{phb}(m_i)$ which has form

$$P_{phb}(m_i) = f \cdot G(m_i - b_1, b_2) + (1 - f) \cdot \left(G(m_i - b_3, b_4) \otimes e^{-b_5 \cdot (m_i - b_3)} \right), \quad (3)$$

where G represents Gaussian function. This part is used to model $B^0 \to J/\psi K^{*0}$ contribution in the fit to extract number of $B_s \to J/\psi\phi$ events.

6.2 Signal mode specifics

To start discussion about specific options of the fit to extract number of $B_s \to J/\psi f_0(980)$ events, it is useful to recall invariant mass distribution shown in Fig. 4. Main feature of that distribution is that in addition to the signal itself, there is a broad structure below masses around 5.26 GeV/c^2. In study of physics backgrounds, this structure was identified as coming from other B^0 and B_s decays. In order to avoid complications arising from parametrization of this structure, we set lower bound of the fit to 5.26 GeV/c^2. Upper boundary of the fit is set to 5.5 GeV/c^2. Second, less prominent but still clear feature is rise of the background between lower boundary of the fit and signal peak itself. While this might be due to real physics (as example $B^0 \to J/\psi\rho$ decay) it is not very clear from data itself that it is real case. The likelihood implemented has two ways of describing resulting distribution. First one is to use polynomial function of relatively high order as only background component. Second is to add Gaussian term to describe this rise and decrease degree of polynomial. In initial tests, both fits yield comparable quality with difference being on 1 σ level. As the second option is more physics motivated we adopt it as default fit and use first option to evaluate systematic uncertainty later on.

Apart of the background, only other specific part is the signal shape. It is parametrized by two Gaussian functions with common mean in form of equation 2. We tested also three Gaussians, but without visible improvement to the fit on the simulated sample and with fraction of third Gaussian being compatible with zero. Parameters are derived from simulated sample or fit to $B_s \to J/\psi\phi$ distribution. The former is used to derive f_i and σ_i parameters while parameters for mean mass M and scaling factor t are caried over from $B_s \to J/\psi\phi$ fit. All signal shape parameters are fixed in the fit of $B_s \to J/\psi f_0(980)$ thus leading to single floating parameter for signal being number of events. The parameters derived from simulation are listed in table 3. In Fig. 9 we show projection of the least square fit used to derive those parameters. In the determination we take into the account specific shape of $f_0(980)$ meson using weighting from section 7. Systematic uncertainty in table 3 reflects uncertainty in the $f_0(980)$ meson mass shape.

Parameter	Value
f_1	$0.823587 \pm 0.0195995 \pm 0.004727$
σ_1 [MeV/c^2]	$9.43652 \pm 0.142973 \pm 0.0335$
σ_2 [MeV/c^2]	$18.3910 \pm 0.464808 \pm 0.36755$

Table 3: Summary of parameters of resolution function/signal shape for $B_s \to J/\psi f_0(980)$ decays derived from simulation. The first uncertainty reflects statistics of the sample while second is systematic uncertainty due to the $f_0(980)$ mass distribution.

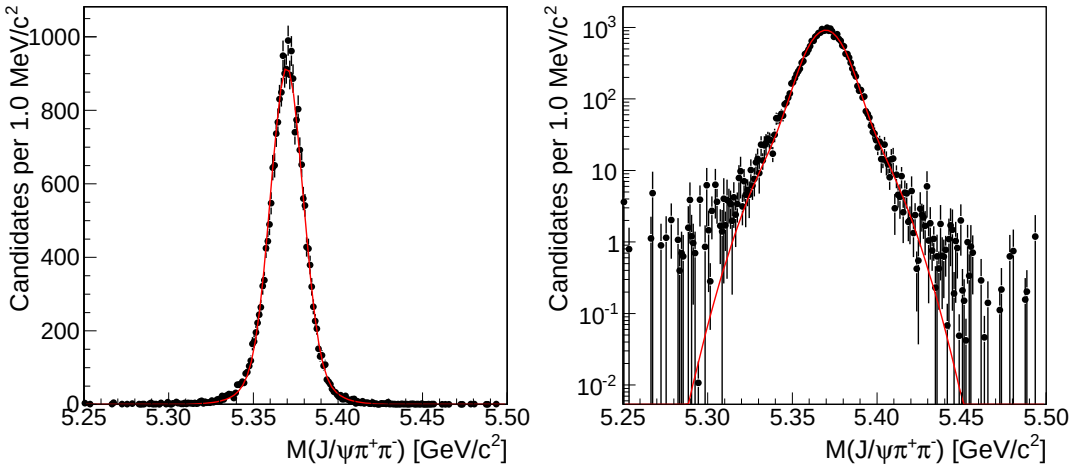

Figure 9: Fit of the resolution function/signal shape for $B_s \to J/\psi f_0(980)$ decays on simulated sample.

6.3 Normalization mode specifics

The fit to extract number of $B_s \to J/\psi\phi$ events has again its specifics related mainly to background. This time two main issues are low amount of combinatorial background and presence of misreconstructed $B^0 \to J/\psi K^{*0}$ decays which peak close to B_s signal. The main issue from having low amount of background is the fact that background parametrization itself can easily pickup imperfections in modeling of signal and bias outcome of the fit. To prevent this issue becoming significant we keep order of polynomial low. Default one being first order polynomial and use variations in order of polynomial to evaluate systematic uncertainties. The fit range is again set to be from 5.26 to 5.5 GeV/c^2.

The misreconstructed $B^0 \to J/\psi K^{*0}$ contribution is parametrized using simulated sample and both shape as well as normalization relative to $B_s \to J/\psi\phi$ is fixed. This contribution is parametrized using equation 3. Summary of its parameters is in table 4 with systematics column given shape parameters for case of phase space $B^0 \to J/\psi K^{*0}$ decay which we use to estimate systematic uncertainty. The distribution itself with pdf overlayed is shown in Fig. 10. The fraction relative to $B_s \to J/\psi\phi$ is given by

$$f_{phb} = \frac{\sigma(B^0)\mathcal{B}(B^0 \to J/\psi K^{*0})}{\sigma(B_s)\mathcal{B}(B_s \to J/\psi\phi)} \cdot \frac{\mathcal{B}(K^* \to K^+\pi^-)}{\mathcal{B}(\phi \to K^+K^-)} \cdot \frac{\epsilon(B_0)}{\epsilon(B_s)}. \quad (4)$$

We take the ratio of B_s and B^0 branching fractions was maesured in Run 1 [13] to be

$$\frac{\sigma(B_s)\mathcal{B}(B_s \to J/\psi\phi)}{\sigma(B^0)\mathcal{B}(B^0 \to J/\psi K^{*0})} = 0.26 \pm 0.08 \pm 0.02 = 0.26 \pm 0.0824. \quad (5)$$

The $\mathcal{B}(K^* \to K^+\pi^-)/\mathcal{B}(\phi \to K^+K^-) = 1.363327 \pm 0.01393995$ is taken from PDG [10]. The ratio of efficiencies is determined from simulated samples. For central value, both B_s and B^0 sample is weighted for best knowledge of polarization and lifetimes and we use pure phase space version to assign uncertainty. Given the large uncertainty in Run 1 measurement for ratio of cross sections times ratio of branching fraction, we use crude estimate of the

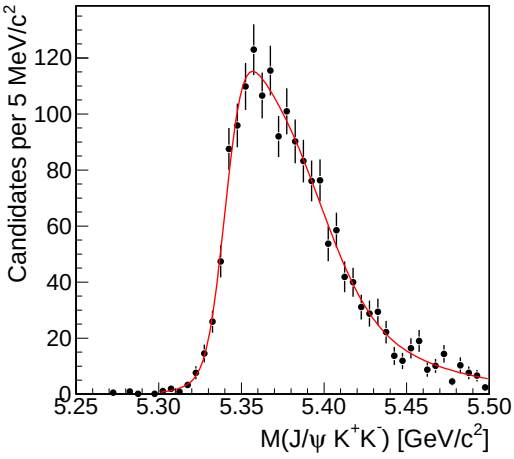

Figure 10: Invariant mass distribution of $B^0 \to J/\psi K^{*0}$ misreconstructed as $B_s \to J/\psi\phi$ derived from simulated data with fit projection overlayed.

Parameter	Value	Systematic
f	0.293674	0.543247
b_1	5.37688	5.35819
b_2	0.0269054	0.0156503
b_3	5.34077	5.37850
b_4	0.00838622	0.00934019
b_5	0.0507128	0.0399945

Table 4: Parameters of pdf to describe $B^0 \to J/\psi K^{*0}$ misreconstructed as $B_s \to J/\psi\phi$.

uncertainty on $\epsilon(B_0)/\epsilon(B_s)$ in which we calculate efficiency ratio weighing only one decay for physics, while leaving other one without weight and taking larger difference as uncertainty. The resulting ratio is $\epsilon(B_0)/\epsilon(B_s) = (5.80228 \pm 0.38379) \cdot 10^{-3}$. Putting all numbers together we obtain $f_{phb} = (3.04 \pm 0.99) \cdot 10^{-2}$.

The signal shape parameters f_i and σ_i are again derived from simulated sample and their values are listed in table 5. Projection of the least square fit used to derive those parameters

Parameter	Value
f_1	0.796128 ± 0.0268839
σ_1 [MeV/c^2]	7.26499 ± 0.119699
σ_2 [MeV/c^2]	13.2821 ± 0.415600

Table 5: Summary of parameters of resolution function/signal shape for $B_s \to J/\psi\phi$ decays derived from simulation.

is in Fig. 11. The simulated sample itself is weighted to take to account discrepancies in p_T distribution, different triggers and physics of the $B_s \to J/\psi\phi$ decays. We also checked that the resolution function does not depend significantly on the momentum distribution and values of physics parameters used.

6.4 Fit Validation

Before performing final fit on data we perform validation using simulated samples. We perform test using three different background parameterizations in configuration suitable for $B_s \to J/\psi f_0(980)$ decay. The three configurations are pure fifth order polynomial, third order polynomial with broad Gaussian function and fifth order polynomial with broad Gaussian. In each case we test three different number of signal events, one corresponding to data itself fit

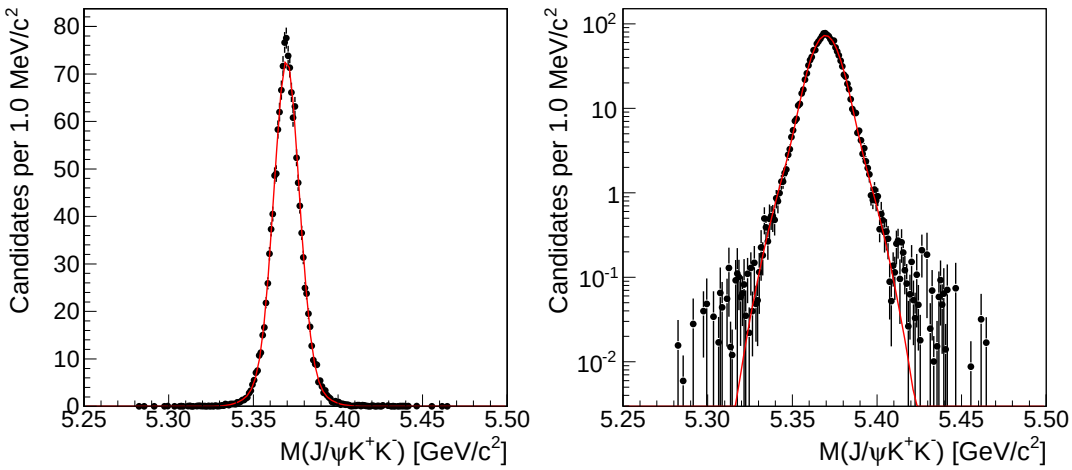

Figure 11: Fit of the resolution function/signal shape for $B_s \to J/\psi\phi$ decays on simulated sample.

with given background function, one with larger number of events and one with lower number of events. We also test configuration used for $B_s \to J/\psi\phi$ decays. We test again three different numbers of signal events while keeping other parameters at their values on data as well as two cases with lower and higher number of background events. Parameters used to simulate experiments are determined from fit to data, except of number of events which were varied. In each sample of simulated experiments we fit pull distribution with Gaussian and report results in table 6. While there are few places with bias which is significant but being below 10% of statistical uncertainty. Given that there is no clear trend in tests we did and size of this bias we neglect it and deal with it in systematic uncertainties.

6.5 Fit results

After fit validation we are prepared to perform final fit. Fit of the $J/\psi K^+K^-$ invariant mass distribution yields 2301.83 ± 49.04 $B_s \to J/\psi\phi$ events. The fit projection and likelihood scan over number of signal events are shown in Fig. 12. In the likelihood profile we perform minimization over other parameters at each point. Running MINOS returns slightly asymmetric uncertainties on number of signal events, namely -48.70 and 49.40. Given the smallness of difference and the fact that MINOS uncertainties are very close to parabolic uncertainties, we keep result simple and use symmetric parabolic uncertainties. Important to note are two other fit parameters, namely mean mass of the B_s signal and scale factor t which scales resolution. Their values are $m(B_s) = 5366.38 \pm 0.207754$ MeV$/c^2$ and $t = 1.11575 \pm 0.0231618$. Those values are used to fix parameters in $B_s \to J/\psi f_0(980)$ fit. Finally in table 7 we show correlation matrix.

In fit to the $J/\psi\pi^+\pi^-$ invariant mass we fix mean B_s mass and resolution scale factor t to the values obtained in $J/\psi K^+K^-$ fit. We use first order polynomial with broad Gaussian as background description as physically this is most sensible option. While increasing order of polynomial by one improves fit slightly, it gives less motivated shape. Additional increase in order of polynomial does not improve fit quality anymore. The fit yields 570.556 ± 37.0679

Background	N	Pull mean $[10^{-2}]$	Pull width	Mean $[10^{-2}]$	RMS
pol 5	541	0.054 ± 1.436	1.000 ± 0.014	-0.189 ± 1.434	1.014 ± 0.010
pol 5	341	1.286 ± 1.410	0.986 ± 0.014	0.836 ± 1.405	0.994 ± 0.010
pol 5	741	-1.139 ± 1.435	1.000 ± 0.014	-0.856 ± 1.424	1.007 ± 0.010
Gauss + p5	540	-1.200 ± 1.814	1.005 ± 0.018	-2.324 ± 1.823	1.023 ± 0.013
Gauss + p5	740	-1.630 ± 1.766	1.016 ± 0.018	-1.733 ± 1.751	1.028 ± 0.012
Gauss + p5	340	-7.146 ± 2.014	1.039 ± 0.020	-6.936 ± 1.996	1.058 ± 0.014
Gauss + p3	505	-4.154 ± 2.073	1.072 ± 0.021	-3.921 ± 2.084	1.096 ± 0.015
Gauss + p3	705	-8.431 ± 2.083	1.042 ± 0.021	-8.647 ± 2.067	1.062 ± 0.015
Gauss + p3	305	-1.310 ± 2.116	1.048 ± 0.021	-0.618 ± 2.147	1.087 ± 0.015
Default	1660.3	-1.071 ± 1.407	0.982 ± 0.014	-0.720 ± 0.014	0.996 ± 0.010
Default	1360.3	-4.560 ± 1.432	0.993 ± 0.014	-5.049 ± 0.014	0.992 ± 0.010
Default	1960.3	-0.906 ± 1.445	1.005 ± 0.014	-0.535 ± 0.014	1.021 ± 0.010
Lower bg	1660.3	-7.260 ± 1.419	0.993 ± 0.014	-1.042 ± 0.014	1.011 ± 0.010
Higher bg	1660.3	-1.782 ± 1.412	0.987 ± 0.014	-2.074 ± 0.014	0.997 ± 0.010

Table 6: Summary of the fitter tests. First section shows configuration used in $B_s \to J/\psi f_0(980)$ while second one is for $B_s \to J/\psi\phi$ fit. Pull mean and width come from Gaussian fit to pull distribution while other two are pure mean and RMS of the pull distributions.

Parameter	$M(B_s)$	N_b	N_s	t	p_1
$M(B_s)$	1.000	-0.003	0.001	0.014	-0.041
N_b	-0.003	1.000	-0.167	-0.222	-0.017
N_s	0.001	-0.167	1.000	0.095	0.007
t	0.014	-0.222	0.095	1.000	0.029
p_1	-0.041	-0.017	0.007	0.029	1.000

Table 7: Correlation matrix of $B_s \to J/\psi\phi$ fit.

$B_s \to J/\psi f_0(980)$ events. The MINOS returns uncertainties -36.8035 and 37.4530. Again as asymmetry is small as well difference from parabolic uncertainty is small, we stick to parabolic uncertainty for the result. The fit projection together with likelihood scan are shown in Fig. 13 and correlation matrix is given in table 8. For number of B^0 events we obtain 179.055 ± 36.0762 events. The value of $-2\ln\mathcal{L}$ is -63018.8. To estimate statistical significance, we repeat fit with number of signal events fixed to zero. For background only fit we obtain $-2\ln\mathcal{L}_0 = -62698.2$. As there is difference of only one parameter between two fits, which is number of signal events, we take square root of difference between two likelihood values as statistical significance of the signal. Numerically it is $\sqrt{-2\Delta\ln\mathcal{L}} = 17.91$. As a sidenote, in background only fit best description is achieved by removing B^0 peak completely. While this would not be best in physics sense, as it gives most conservative significance we use this fit. If we would force fit to describe B^0 peak, significance of the signal would be larger.

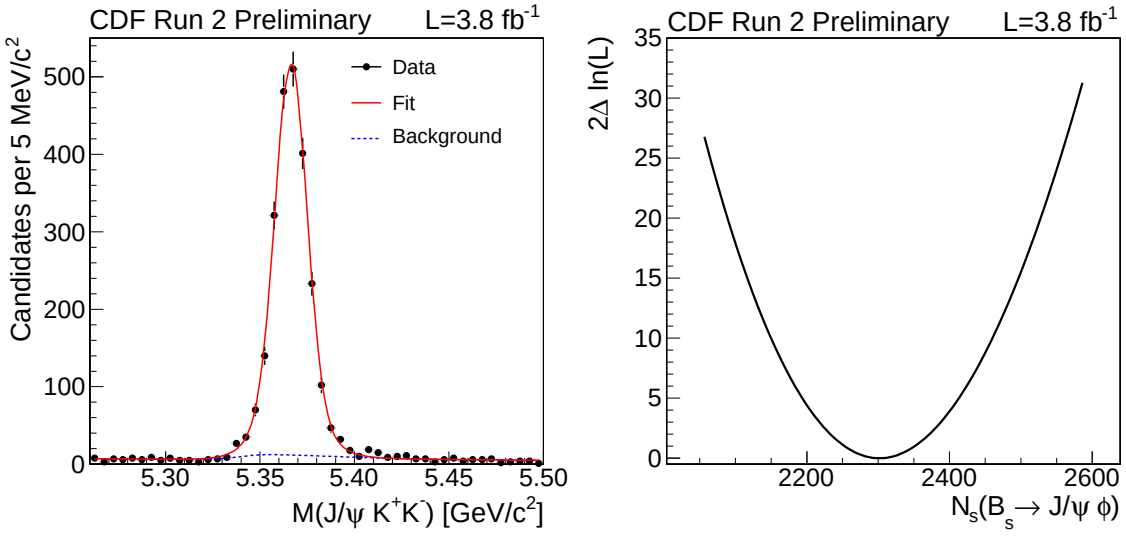

Figure 12: Fit projection of the final fit on $B_s \to J/\psi\phi$ decay mode (left) and likelihood profile of number of signal events (right).

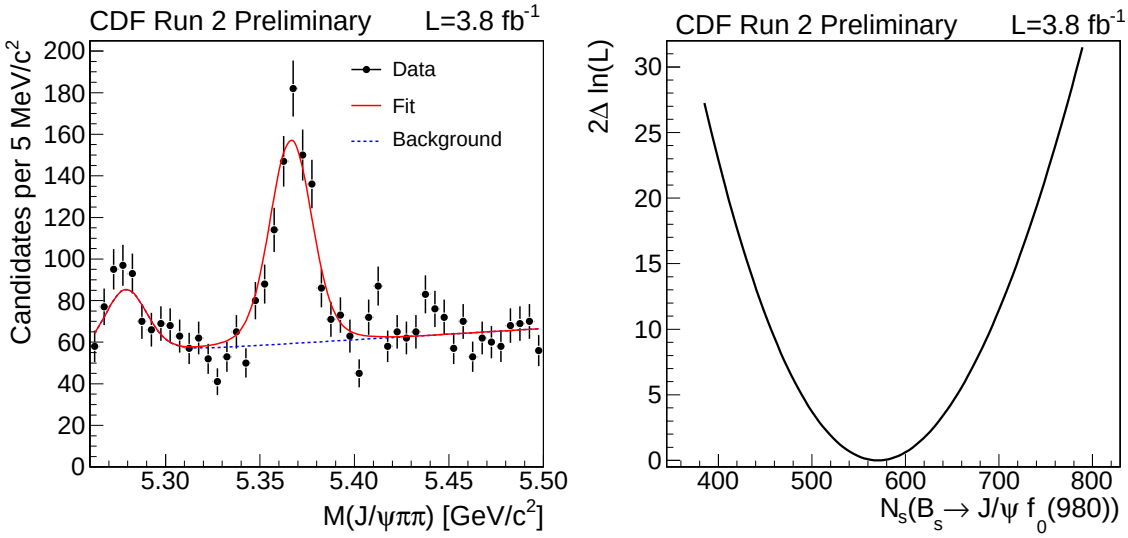

Figure 13: Fit projection of the final fit on $B_s \to J/\psi f_0(980)$ decay mode (left) and likelihood scan of number of signal events (right).

Parameter	N_{B^0}	N_b	N_s	p_1
N_{B^0}	1.000	-0.548	0.260	0.658
N_b	-0.548	1.000	-0.418	-0.442
N_s	0.260	-0.418	1.000	0.245
p_1	0.658	-0.442	0.245	1.000

Table 8: Correlation matrix of $B_s \to J/\psi f_0(980)$ fit.

7 Relative efficiency

Our aim is to measure relative branching fraction between decays $B_s \to J/\psi f_0(980)$ and $B_s \to J/\psi\phi$. This can be obtained using relation

$$\frac{\mathcal{B}(B_s \to J/\psi f_0(980))}{\mathcal{B}(B_s \to J/\psi\phi)} \frac{\mathcal{B}(f_0(980) \to \pi^+\pi^-)}{\mathcal{B}(\phi \to K^+K^-)} = \frac{N(B_s \to J/\psi f_0(980))}{N(B_s \to J/\psi\phi)} \frac{\epsilon(B_s \to J/\psi\phi)}{\epsilon(B_s \to J/\psi f_0(980))} \quad (6)$$

where $N(B_s \to J/\psi f_0(980))$ and $N(B_s \to J/\psi\phi)$ are number of events measured in section 6.5 and $\epsilon(B_s \to J/\psi f_0(980))$ with $\epsilon(B_s \to J/\psi\phi)$ are efficiencies to reconstruct and select candidate of given decay. At this stage we factor out hadron branching fractions as they are entered from external source, rather then being experiment specific. In this section we evaluate $\epsilon_{rel} = \epsilon(B_s \to J/\psi\phi)/\epsilon(B_s \to J/\psi f_0(980))$.

To estimate ϵ_{rel} we generate independent sample using 6.1.4mc.m release using patch s tarball after merge of B specific setup to default high pt tarball. In this sample we force with equal probabilities B_s to decay to one of the two decays needed. Sample is generated with single lifetime of $c\tau = 441\mu$m, phase space decay model for all particles involved and default mass model for $f_0(980)$. During calculation of the ϵ_{rel} we weigh sample to appropriate physics distributions. For $B_s \to J/\psi\phi$ this involves weighting using prescription in section 2.1. For $B_s \to J/\psi f_0(980)$ we reweigh helicity distribution of the positive muon to $\sin^2\theta$ distribution. As it is CP-odd final state its lifetime is different from lifetime distribution of $B_s \to J/\psi\phi$. We assume no CP violation as good approximation of standard model and use lifetime corresponding to heavy mass eigenstate. Numerical value follows directly from parameters used to weigh $B_s \to J/\psi\phi$ decay. Last piece needed is mass distribution of $f_0(980)$. While EvtGen uses usual Breit-Wigner distribution to describe it, experiments dealing with $f_0(980)$ typicallt use Flatté distribution [14]. It is defined by amplitude

$$A(m) = \frac{1}{m_0^2 - m^2 - im_0\Gamma} \quad (7)$$

with m being actual true mass of the candidate, m_0 being central mass and Γ being

$$\Gamma = g_\pi\sqrt{m^2/4 - m_\pi^2} + g_K\sqrt{m^2/4 - m_K^2} \quad (8)$$

above threshold for two kaons and

$$\Gamma = g_\pi\sqrt{m^2/4 - m_\pi^2} + ig_K\sqrt{m_K^2 - m^2/4} \quad (9)$$

below threshold. In this expression, g_π and g_K are two coupling constants, m_π and m_K masses of charged pions and kaons. Invariant mass distribution is than described by the amplitude squared. Numerical values are taken from BES measurement [15] to be

$$\begin{aligned} m_0 &= 965 \pm 8 \pm 6 \text{ MeV}/c^2, \\ g_\pi &= 165 \pm 10 \pm 15 \text{ MeV}/c^2, \\ g_K/g_\pi &= 4.21 \pm 0.25 \pm 0.21. \end{aligned}$$

After weighting simulated sample for best physics knowledge we obtain relative efficiency $\epsilon_{rel} = 1.17775 \pm 0.0124304$ where uncertainty reflects statistics of the simulated sample.

Relatively large difference from unity is caused mainly by the p_T cut on the ϕ and $f_0(980)$ tracks. In case of ϕ decay, kaons have little momenta in ϕ rest frame and thus when boosted they are close to collinear with momentum along ϕ direction. In contrast, pions in $f_0(980)$ decay have larger momentum and thus larger chance that one pion will get small p_T when boosted to the laboratory frame.

8 Systematic uncertainties

There are several sources of systematic uncertainties, which affect result presented in this note. They can be divided into two main classes, first one having effect throught the extraction of number of events and second one through efficiency. When talking about significance of the signal, only first class plays role. Given the high statistics of the signal for first class of uncertainties, we will refit data with modified assumptions and take difference to the default fit as systematic uncertainty.

8.1 Background parametrization in $B_s \to J/\psi\phi$

The combinatorial background we use is purely phenomenological rather than fixed by some underlying theory. Therefore it is not garantied that parametrization we use is correct one. To asses how much result depends on the exact parametrization of the combinatorial background, we repeat fit with uniform combinatorial background and using second order polynomial. We calculate difference to the default fit and take larger of the two differences as systematic uncertainty. For N_s it is 34.35, for $M(B_s)$ its size is 0.01 MeV/c^2 and for the resolution scale factor t its size is 0.0208.

One thing to note here is the issue of extremaly small combinatorial background. With such small background and large signal, it is very difficult to disantangle imperfections in modeling of the background and tails of the resolution function. From this reason, even if increasing order of polynomial function increases quality of the fit, we do not go above second order.

8.2 Physics background in $B_s \to J/\psi\phi$

In the fit to $J/\psi K^+K^-$ invariant mass we fix amount of the misreconstructed $B^0 \to J/\psi K^{*0}$ as well as its shape. To estimate systematic uncertainty due to the size of $B^0 \to J/\psi K^{*0}$ contribution, we vary amount by one standard deviation and take larger difference to the central fit as uncertainty. In numerical value it gives uncertainty of 13.13 $B_s \to J/\psi\phi$ events. To estimate systematic uncertainties due to limited knowledge of the shape of physics background, we derive alternative shape using phase space decay for $B^0 \to J/\psi K^{*0}$ and refit data with new shape. We find difference of 1.06 events which we assign as systematic uncertainty. adding two contributions in quadrature we obtain uncertainty of 13.17 due to limited knowledge of physics background.

In the same time we also check variations of B_s mass and resolution scale t as they are used in $B_s \to J/\psi f_0(980)$ fit. From variations we assign uncertainties of 0.0 MeV/c^2 on $M(B_s)$ and $7.7 \cdot 10^{-3}$ on resolution scale t.

8.3 Resolution in $B_s \to J/\psi\phi$

The resolution function we choose is rather arbitrarily composed of two Gaussian. In addition being derived from simulation, limited knowledge of physics of the B_s production and decay to $J/\psi\phi$ final state enters. To estimate uncertainty, we perform two alternative fits.

In first one we use same physics knowledge as in default resolution function, but use only single Gaussian rather than three to parametrized resolution function. Taking difference between default and modified fit yields uncertainty of 31.58 events.

In second alternative fit we remove weighting of the events, which modifies both production p_T spectrum as well as uses phase space decay distribution. Resolution function is parametrized with three Gaussians in this case. From this we assign uncertainty of 2.27 events.

The two effects are basically independent and therefore we add both contributions in quadrature which gives uncertainty of 31.58. In same time we also track changes in central position of the signal and resolution scale t. Uncertainty on $M(B_s)$ is 0.01 MeV/c^2 and on t is 0.02462.

8.4 Background parametrization in $B_s \to J/\psi f_0(980)$

Similar as before, parametrization of the combinatorial background is purely phenomenological. In order to test sensitivity to the parametrization we refit data using uniform distribution and parabola instead of linear function. In both cases B^0 peak is untouched with normalization floating in the fit. The larger yield difference is assigned as systematic uncertainty. In numbers it equals 21.738 events.

8.5 Resolution in $B_s \to J/\psi f_0(980)$

Similar to the case of $B_s \to J/\psi\phi$, also here we test several effects. First one is the use of single Gaussian instead of double Gaussian for resolution function. This contributes with 4.26 events uncertainty. Second effect is question of production momentum dependence. Again removing weighting which yields rather poor agreement with data and parametrizing resolution function with double Gaussian yields uncertainty of 8.82 events. Finally as we fix resolution scale factor t in the fit, we redo fit with values larger and smaller by one sigma (including systematic variation) and take larger of two difference as uncertainty. This adds another 7.48 events uncertainty. Summing all three contributions in quadrature yields uncertainty of 12.324 events.

8.6 Fixed B_s mass

Mass of the B_s is fixed in the fit of $B_s \to J/\psi f_0(980)$ decay to the value found in $B_s \to J/\psi\phi$ fit. We vary its value within statistical and systematical uncertainty and find uncertainty of 2.845 events.

8.7 Effect on significance

All systematic variations we investigate can be separated to two classes. One class affects shape and position of the signal, while other contains changes in background description. During refits to estimate systematic uncertainties we tracked also changes to the value of

likelihood function at minimum. The largest change for first class gives decrease of $-2\Delta \ln \mathcal{L}$ of 5.5, which provides insignificant change to the significance derived from $-2\Delta \ln \mathcal{L} = 320.6$. It can be understood by fact that simpler resolution function provides worst description as it clearly cannot describe all details of the detector response. While combination of several modifications can yield larger change, it is unlikely that it would decrease significance considerably. For different background parametrizations we redo two fits with and without signal and find that smallest $-2\Delta \ln \mathcal{L}$ is obtained using using third order polynomial. With value of $-2\Delta \ln \mathcal{L} = 271.1$ which would decrease significance to 16.47σ. As this is only small change compared to significance itself we conclude that signal is solid under systematic variation and rather then trying to pinpoint significance including systematic effects we quote just statistical significance of the signal.

8.8 B_s momentum distribution

Production momentum distribution defines momenta of all daughter particles and as we place requirement on p_T of the tracks as well as p_T of the B_s candidate its imperfect knowledge can have effect on relative efficiency ϵ_{rel}. While effect should mostly cancel in ratio, there can be some second order effect. To estimate it, we recalculate efficiency without p_T reweighting and take difference to the default ϵ_{rel} as uncertainty. Numerically uncertainty due to this source is 0.01067.

8.9 Physics of the $B_s \to J/\psi\phi$ and $B_s \to J/\psi f_0(980)$ decays

As the physics behind two studied decays is different with some quantities which are determined by the experiment, we expect some effect on the relative efficiency. Given the difference we even don't expect that effects would cancel out by default. Effects we investigate here are limited knowledge of the mass shape of $f_0(980)$ resonance, B_s lifetime and decay width difference and polarization amplitudes in $B_s \to J/\psi\phi$ decays.

The mass shape of the $f_0(980)$ resonance is strongly affected by its proximity to the $K\bar{K}$ threshold which modifies it from usual Breit-Wigner shape. As mentioned during efficiency evaluation, we use shape derived from the BES experiment. In order to estimate our sensitivity to the shape of $f_0(980)$ we vary parameters withing uncertainties, with exception of mean mass which we vary from 965 MeV/c^2 to 980 MeV/c^2. Uncertainty we obtain is 0.02470. It is mainly due to the fact that differences in mass shape cause small modifications of the p_T distributions of pions from decay of $f_0(980)$ and thus p_T requirement has slightly different efficiency.

One of the peculiarity of the B_s system is non-zero decay width difference. Assuming no CP-violation, which is good assumption in the standard model, this translates to different lifetime of CP-odd and CP-even final states. As $B_s \to J/\psi f_0(980)$ is CP-odd and $B_s \to J/\psi\phi$ is mixture with dominant CP-even contribution, we expect effect proportional to the knowledge of decay width difference. To estimate systematic uncertainty, we calculate efficiency in all combinations where lifetime and decay width can be one sigma higher, measured central value or one sigma lower and take largest deviation from default relative efficiency. The obtained uncertainty is 0.01981.

Other part of the experimental input is in polarization amplitudes in $B_s \to J/\psi\phi$ decay including their relative phase.Those enter the efficiency in two ways. First they define angular

Effect	$B_s \to J/\psi\phi$	$B_s \to J/\psi f_0(980)$	ϵ_{rel}
Combinatorial bg.	34.35	21.738	-
Physics bg	13.17	-	-
Resolution	31.58	9.79	-
Resolution scale	-	7.48	-
B_s mass	-	0.472	-
Yield summary	48.48	25.15	-
MC statistics	-	-	0.0124304
Momentum distribution	-	-	0.01067
Physics of decays	-	-	0.0328905
Trigger paths	-	-	0.01636
Efficiency summary	-	-	0.04022187

Table 9: Summary of assigned systematic uncertainties.

distributions which determine momentum distributions of tracks. Second they define relative fraction between CP-even and CP-odd component and reflects through lifetime and decay width difference. On the other hand, they are measured more precisely then decay width difference, so they overal effect is expected to be smaller. To estimate systematic uncertainty we vary each of the three parameters (A_0, $A_{||}$ and $\phi_{||}$) by one sigma around central value and add three differences in quadrature. We obtain uncertainty of 0.0089028. Given the size of this contribution compared to previous one we don't expect that more thourough evaluation would cause significant difference in total uncertainty.

We consider three effects to be more less independent and therefore sum them in quadrature to obtain systematic uncertainty due to the modeling of $B_s \to J/\psi f_0(980)$ and $B_s \to J/\psi\phi$ decays. It contributes to the systematic uncertainty by 0.0328905.

8.10 Trigger paths differences

In the analysis we don't separate between different trigger paths when calculating branching fraction. It is therefore important that the mixture of different trigger pathes in simulation is correct unless all trigger paths have same relative efficiency. When we do selection, we select only handful of classes, with similar requirements for all trigger paths in same class. During efficiency calculation we reweigh fraction of each class to match one in the data, so in first order our average relative efficiency is correct. On the other hand, it is practically impossible to decide to which class event should belong and thus our observation in data does not have to precisely reflect true. To estimate uncertainty, we calculate relative efficiency for each class and take half of the difference between largest and smallest as uncertainty. With efficiencies of 1.17064 for class 1, 1.21236 for class 2 and 1.17095 for class three we assign uncertainty of 0.01636 coming from trigger mixture of our sample.

8.11 Summary

In table 9 we summarize all systematic uncertainties assigned. They fall in two broad classes, one which affects yields and other affecting relative efficiency. We keep those separate at this stage and propage them in final result. All contributions to each class are treated as

Parameter	Fit 1	Fit 2
m_0	989.6 ± 9.9	1003.3 ± 5.8
g_π	141.4 ± 18.8	141.0 ± 15.7
g_K/g_π	2.32 ± 1.33	4.21 (Fixed)
χ^2/ndf	36.34/31	37.53/32
Fit probability	23.38%	23.05%

Table 10: Parameters for dipion invariant mass distribution fit using Flattè distribution.

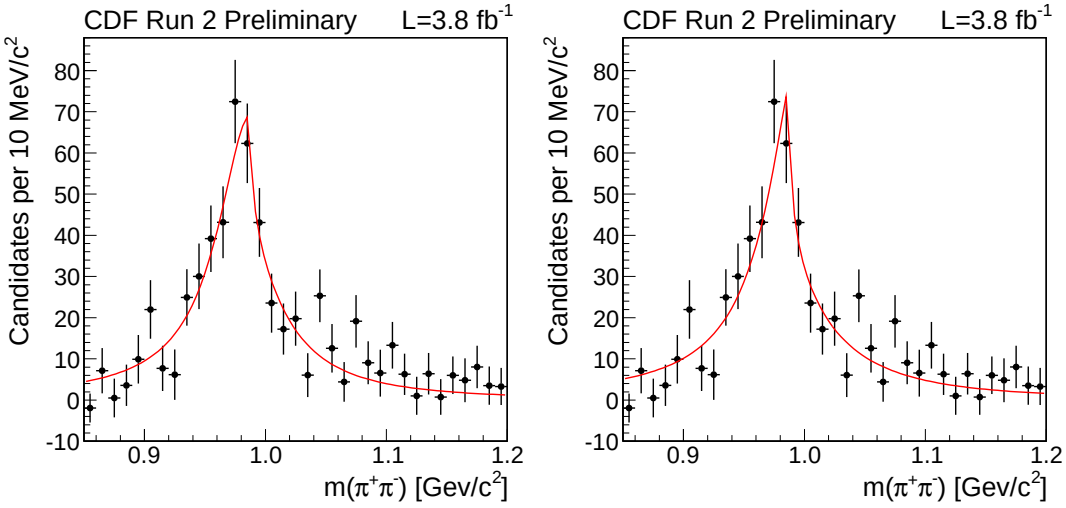

Figure 14: The dipion invariant mass distribution after sideband subtraction with fit projection overlayed. Fit uses Flattè distribution. On left side all parameters are floating and on right side g_K/g_π is fixed to BES value.

independent and added in quadrature to obtain total uncertainty. For $B_s \to J/\psi f_0(980)$ yield the systematic uncertainty is 25.15 events, for $B_s \to J/\psi\phi$ yield it is 48.48 events and for ϵ_{rel} we obtain uncertainty of 0.04022187.

9 Nature of the signal

In this section we perform basic checks to show that signal we interprete as $B_s \to J/\psi f_0(980)$ decay is consistent with this interpretation. Two main features are exploited, invariant mass of the two pions and helicity angle of the positive pion and muon. If our interpretation is correct than the dipion invariant mass should be consistent with $f_0(980)$ mass shape and the distribution of helicity angle after correcting for efficiency should be uniform. The muon helicity angle should exhibit $1 - cos^2(\theta_{\mu^+})$ behaviour. To extract distributions for pure signal we fit $J/\psi\pi^+\pi^-$ mass distribution in bins of given variable. As the background shape depends on the pion helicity angle we restrict fit range from 5.26 GeV/c^2 to 5.45 GeV/c^2. This avoids large changes in the background shape and makes fits on rather low statistics more stable.

We show dipion invariant mass distribution in Fig. 14. We fit distribution using Flattè formula and list obtained values in table 10. As we have very limitted sensitivity to g_K/g_π we

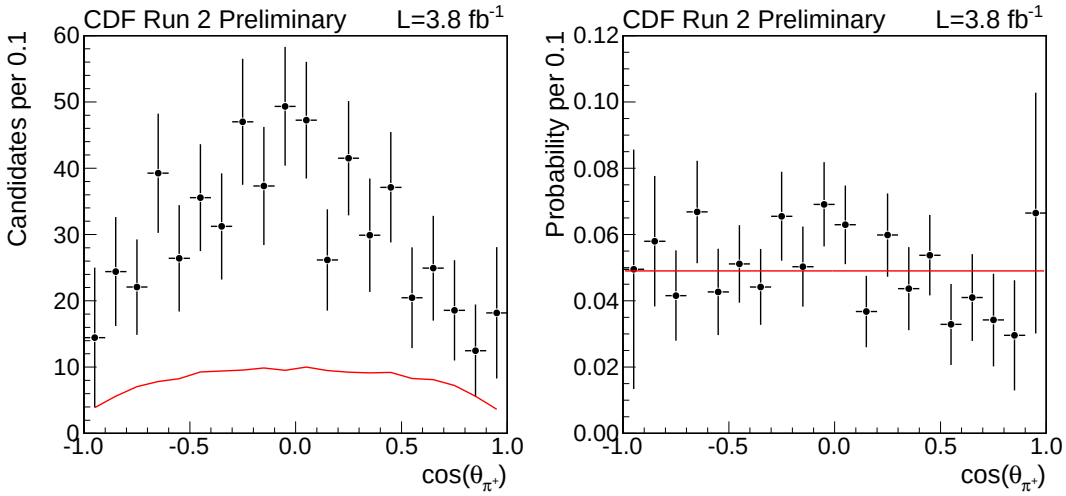

Figure 15: The distribution of helicity angle of positive pion for B_s signal events before correction for acceptance (left) and after correction (right). The red line in left show arbitrarily normalized acceptance.

also repeat fit where we fix $g_K/g_\pi = 4.21$, value measured by BES. While agreement between BES and us is not perfect it is of acceptable quality given the general disagreement between different measurements of $f_0(980)$ shape parameters [10].

The distribution of helicity angle of positive pion is shown in Fig. 15. We use simulation used to calculate efficiency to derive acceptance correction and correct distribution from data. After correction, except of bins on the edges, the distribution is very well consistent with uniform distribution. In order to quantify consistency with uniform distribution we fit it with constant and obtain $\chi^2/ndf = 14.89/19$. As distribution is expected to be symmetric we fit also distribution of absolute value of $\cos(\theta_{\mu^+})$. The fit is shown in Fig. 17 and has $\chi^2/ndf = 10.39/9$. In Fig. 16 we show positive muon helicity angle. Fit with theoretical distribution gives $\chi^2/ndf = 7.22/19$. The fit to distribution of absolute values shown if Fig. 17 gives $\chi^2/ndf = 3.65/9$.

Overal from those check we can conclude that $B_s \to J/\psi f_0(980)$ signal we observe in our data is consistent with expactations for this decay.

10 Result

At this stage we have all necessary inputs available and can proceed to calculate final result. Summary of needed inputs to equation 6 is in table 11. As a main result we obtain

$$R = \frac{\mathcal{B}(B_s \to J/\psi f_0(980))}{\mathcal{B}(B_s \to J/\psi\phi)} \frac{\mathcal{B}(f_0(980) \to \pi^+\pi^-)}{\mathcal{B}(\phi \to K^+K^-)} = 0.291645 \pm 0.0199 \pm 0.0173839. \quad (10)$$

Using world average values for ϕ and $B_s \to J/\psi\phi$ branching fractions we obtain

$$\mathcal{B}(B_s \to J/\psi f_0(980))\mathcal{B}(f_0(980) \to \pi^+\pi^-) = (1.854 \pm 0.127 \pm 0.111 \pm 0.570) \cdot 10^{-4}, \quad (11)$$

where first uncertainty is statistical, second systematical and third one due to uncertainty on branching fractions. We do not calculate separate branching fraction of pure $B_s \to J/\psi f_0(980)$

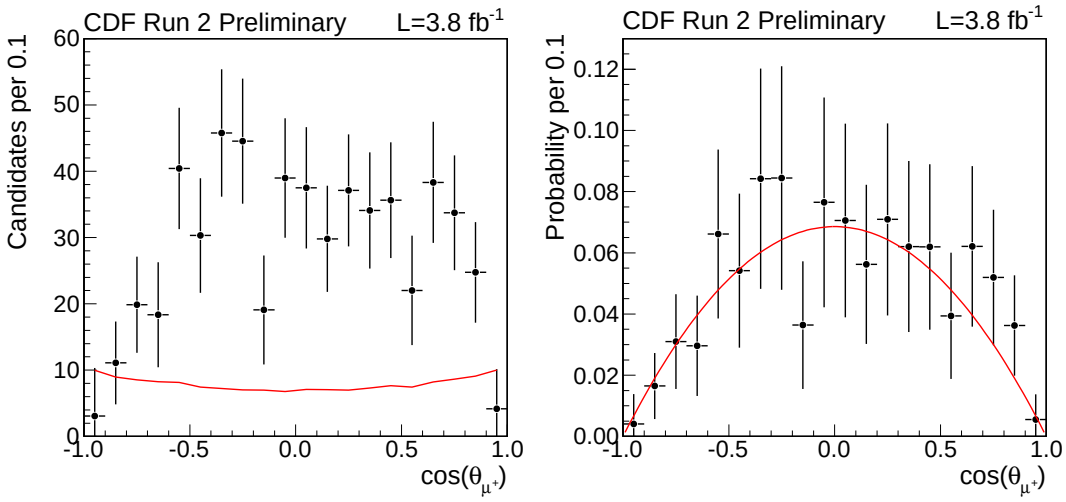

Figure 16: The distribution of helicity angle of positive muon for B_s signal events before correction for acceptance (left) and after correction (right). The red line in left show arbitrarily normalized acceptance.

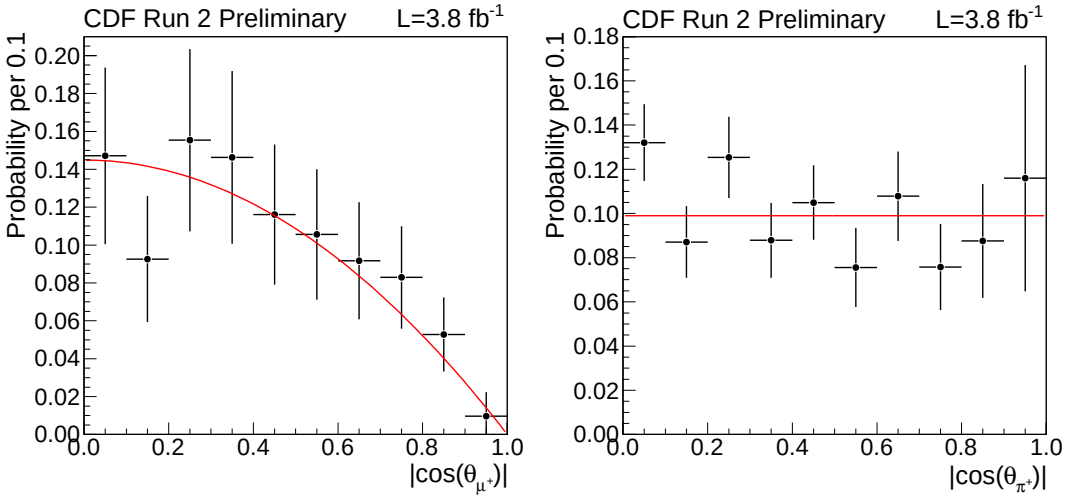

Figure 17: The distribution of helicity angle of positive muon (left) and positive pion (right) for B_s signal events after correction.

Quantity	Value
$N(B_s \to J/\psi f_0(980))$	$570.556 \pm 37.0679 \pm 25.15$
$N(B_s \to J/\psi\phi)$	$2301.83 \pm 49.04 \pm 48.48$
$\epsilon_{rel} = \epsilon(B_s \to J/\psi\phi)/\epsilon(B_s \to J/\psi f_0(980))$	1.17775 ± 0.04022187
$\mathcal{B}(\phi \to K^+K^-)$	0.489 ± 0.005
$\mathcal{B}(B_s \to J/\psi\phi)$	$(1.3 \pm 0.4) \cdot 10^{-3}$

Table 11: Summary of inputs to calculation of branching fraction for $B_s \to J/\psi f_0(980)$ decay.

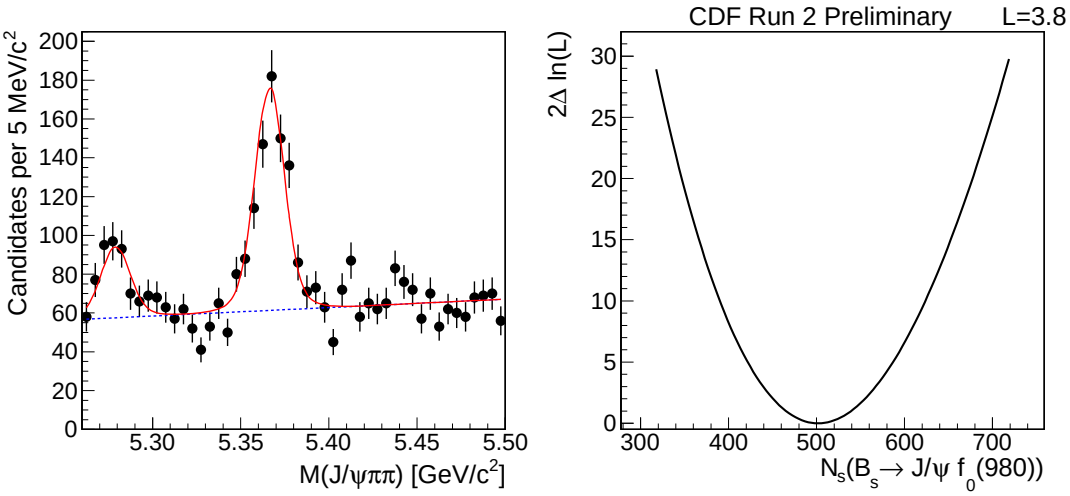

Figure 18: Fit projection of the fit with scale factor t floating on $B_s \to J/\psi f_0(980)$ decay mode (left) and likelihood scan of number of signal events (right).

decay due to the purely known $f_0(980)$ branching fractions. The value of R is consistent with expectations which are in the range of 0.2 - 0.5.

While we interprete our signal as being due to the $B_s \to J/\psi f_0(980)$ decay and support this by study in previous section, there is small possibility that signal is contaminated by other f resonances. As we use wide $\pi^+\pi^-$ window, we cannot exclude that there is some contribution from $f_2(1270)$ or $f_0(1370)$, but as in our reconstruction we stopped at $m(\pi^+\pi^-) = 1.2$ GeV/c^2 we cannot investigate this further. In addition to observing signal, we measure relative branching fraction, which is in the expected range. This observation provides interesting possibility of using pure CP-eigenstate to measure lifetime or CP violation in B_s system.

11 Changes for paper

In this section we describe changes to the analysis and results triggered by the work on the lifetime in $B_s \to J/\psi f_0(980)$ decays. There it was observed that fit in which mass resolution scale factor is left floating gives significantly better description of data compared to fit with mass resolution scale factor fixed to the value obtained on $B_s \to J/\psi\phi$ decays. Trigger by this after discussion with godparents we decided to run analysis with mass resolution scale factor floating also on $B_s \to J/\psi f_0(980)$ fit.

As a result of leaving scale factor t floating we find number of $B_s \to J/\psi f_0(980)$ events to be $N = 501.743 \pm 36.7852$ with MINOS uncertainties being close to parabolic ones and asymmetry is on the same level as before. In table 12 we show correlation matrix of the fit. The number of $B^0 \to J/\psi\pi^+\pi^-$ events return by the fit is 160 ± 30 events.

Given the change in the fit we also evaluate systematic uncertainties on the $B_s \to J/\psi f_0(980)$ yield using new fit. The procedure is same as before, we just drop variation of scale factor t which is now free fit parameter and thus its uncertainty is already included in the statistical uncertainty. The systematic uncertainty on due to the assumed background shape is found to be 16.259 events. Uncertainty due to resolution has two contributions,

Parameter	N_{B^0}	N_b	N_s	p_1	t
N_{B^0}	1.000	-0.476	0.253	0.233	0.574
N_b	-0.476	1.000	-0.428	-0.327	-0.358
N_s	0.253	-0.428	1.000	0.448	0.228
p_1	0.233	-0.327	0.448	1.000	0.194
t	0.574	-0.358	0.228	0.194	1.000

Table 12: Correlation matrix of $B_s \to J/\psi f_0(980)$ fit with scale factor t floating.

first from changing double Gaussian to single Gaussian (7.537 events). Second one is from modifiying kinematical distribution of simulated events which reflects to different ratio between areas of two Gaussians as well as for widths itself (2.364 events). In total, imperfect knowledge of resolution contributes by 7.8971 events. The fixed B_s mass contributes to systematic uncertainties by 0.1203 events. Summing all contributions in quadrature we obtain systematic uncertainty on the yield of 18.075778 events.

For completetness we obtain width scale $t = 0.829 \pm 0.054$ which is significantly smaller than one obtain in $B_s \to J/\psi\phi$. While the large difference in scale factors is somehow surprising, it is probably not unexpected. We would like to remind reader, that already in study of $\psi(2S) \to J/\psi\pi^+\pi^-$ decays we observed scale factor of 1.052, which is bellow value seen in $B_s \to J/\psi\phi$ decays. As in reconstruction we constrain in all cases dimuons to the world average J/ψ mass the two tracks from $f_0(980)$ or ϕ are responsible for most of the resolution. But the decays $f_0(980) \to \pi^+\pi^-$ and $\phi \to K^+K^-$ are rather different. While in former each pion has significant momentum in $f_0(980)$ rest frame, kaons from ϕ decay are almost in rest. After transformation to the laboratory system, pions can gain large momentum imbalance and will have relatively large opening angle. On contrary two kaons will be almost collinear. In addition, $B_s \to J/\psi f_0(980)$ vertex is expected to be measured more precisely than $B_s \to J/\psi\phi$, where kaons contribution is smaller. When measuring invariant mass, decay $B_s \to J/\psi f_0(980)$ will depent mostly on how well we can measure absolute momentum, while $B_s \to J/\psi\phi$ is more sensitive to opening angle between two kaons. While absolute momentum is dominated more by COT, angle is much more sensitive to silicon detector and thus two decays probe different aspects of our detector and simulation.

Propagating new $B_s \to J/\psi f_0(980)$ yield and uncertainties through formulas leads to

$$R_{f_0/\phi} = 0.256721 \pm 0.0196 \pm 0.01384 \tag{12}$$

and

$$\mathcal{B}(B_s \to J/\psi f_0(980))\mathcal{B}(f_0(980) \to \pi^+\pi^-) = (1.63198 \pm 0.12460 \pm 0.08799 \pm 0.5024) \cdot 10^{-4} \tag{13}$$

where last uncertainty is due to the uncertainties on world average branching fractions.

A What has change

A.1 Between version 1.0 and 2.0

1. While attempting to understand bump at lower masses I could convince myself it is real B^0 signal. This resulted in change of fit function with broad Gaussian removed and

another signal structure added. Additional signal has same shape as $B_s \to J/\psi f_0(980)$ and its position is fixed to 5.279 GeV/c^2.

2. For better consistency between two channels change resolution function of $B_s \to J/\psi\phi$ to double Gaussian as is used for $B_s \to J/\psi f_0(980)$.

3. As a result of number 2 alternative model for systematic uncertainty on signal shape for $B_s \to J/\psi\phi$ is single Gaussian and double Gaussian with modified momentum distribution and phase-space decay.

4. Some readers might note that systematic uncertainty on $B_s \to J/\psi\phi$ increased significantly. This comes from the fact that third Gaussian had only tiny fraction and thus modification from three to two Gaussians was practically tiny. Now modification from two to one Gaussian is much larger and thus uncertainty increased.

5. As a result of change in background description for $B_s \to J/\psi f_0(980)$ in systematic uncertainty evaluation we change to parabola and keep B^0 signal always in the fit unchanged.

6. Change lower fit limit from 5.25 GeV/c^2 to 5.26 GeV/c^2, motivated by simplification based on simulation.

7. As a result of changes to fits, all systematic uncertainties on yields are re-evaluated. This results in small numerical changes in many place, but none of them is really incompatible with previous version.

8. Bug was found in sideband subtruction for helicity angles. After fixing, strong mass dependence was observed for pion helicity angle and thus subtraction using only upper sideband didn't work anymore. Changed to slicing, where I fit $J/\psi\pi^+\pi^-$ spectrum in bins of given variable. This is done for all distributions in chapter 9.

9. Added also muon helicity angle distribution for signal.

10. Found bug in evaluation of systematic uncertainty due to lifetime and decay width difference. Fixing it significantly decreased this uncertainty.

A.2 Between version 2.0 and 3.0

1. Fixed bug in calculation of lifetime of heavy and light eigenstate. Previously they were swaped. New version is using

$$\begin{aligned} \Gamma_H &= 1./\tau - 0.5 \cdot \Delta\Gamma \\ \Gamma_L &= 1./\tau + 0.5 \cdot \Delta\Gamma \end{aligned}$$

A.3 Between version 3.0 and 4.0

Leave mass resolution scale factor floating also on $B_s \to J/\psi f_0(980)$ fit which gives better description of data. All changed numbers are documented in separate section.

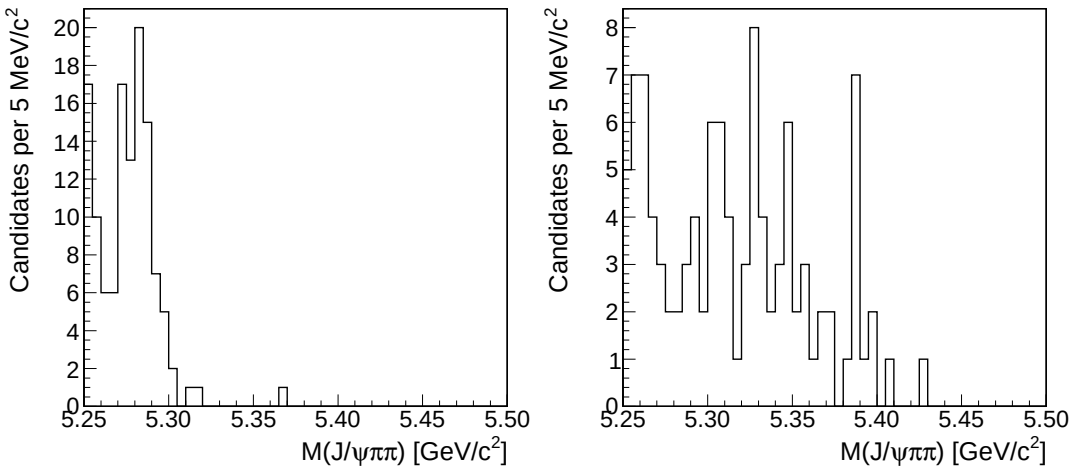

Figure 19: The invariant mass distribution of $J/\psi\pi^+\pi^-$ candidates in simulated events for B^0 decays (left) and B_s decays (right).

B Why to change fit model

Earlier versions of this note used different background parametrization. At the time of pre-blessing, background was described by a linear function together with broad Gaussian which was centered around 5.27 GeV/c^2. During pre-blessing question came up whether we understand structure around B^0 mass, close to the lower bound of the fit. In what follows we describe our way to understanding of the structure, which motivated change in the background description in the analysis.

From long CDF experience it is rather clear that structure we try to understand is due to real b-hadron decays rather than from some other source. As we have inclusive J/ψ simulation in hands, we had a closer look to it to try to answer question whether we can motivate broad Gaussian in our fit function from it. In Fig. 19 we show once again simulated events from Figs. 5 and 7, but this time using same B_s mass range and binning as our final plots. Inspecting them on this scale provides two pieces of information. First one is that B_s decays could possible look as a broad Gaussian. Second part of the information is that B^0 distribution looks in first few bins quite similar to real data. In addition, relative normalization of the two distribution is such that B_s distribution should be scaled down by factor 4.7 before adding to the B^0 contribution. The factor 4.7 takes into account relative production fraction for B_s and B^0 mesons as well as different amount of decays simulated. This means that while B_s decay can justify broad Gaussian, its contribution is much smaller than B^0 contribution and if we would claim we describe B_s part we would first need to deal with B^0.

Such argument immediately suggest that we might in fact deal with real B^0 and the question arises whether we can in some way check that it is a case. First idea is to put B^0 signal in the fit and see whether we can describe data. As a quick check we took our original fitter and reused broad Gaussian for the test. In test itself we fixed position of the Gaussian to world average B^0 mass and width to 12 MeV/c^2. Projection of this fit is shown in Fig. 20. As we can see distribution is described rather well (only small change in likelihood value compared to fit with broader structure). Only part which is not properly described are first

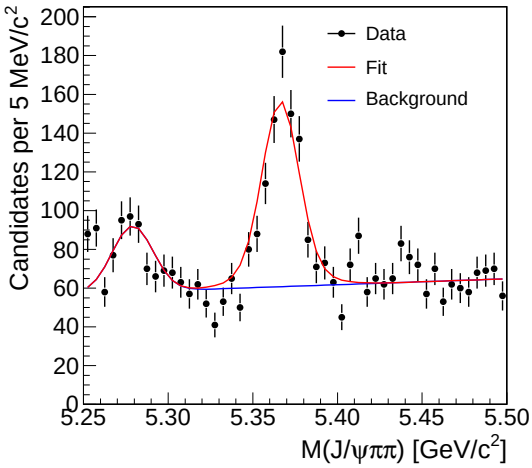

Figure 20: The invariant mass distribution of $J/\psi\pi^+\pi^-$ candidates with fit projection using linear combinatorial background and single Gaussian for B^0 background.

two bins, where we expect start of the rise of mis-reconstructed events. In this fit we find about 200 events with roughly 15-20% uncertainties.

In section 5 we identified one source of background, which gives correctly reconstructed B^0, which is decay $B^0 \to J/\psi\rho^0$. From known branching fraction and our $B_s \to J/\psi\phi$ signal we can derive expectation for number of events we should see in our spectrum. This expectation is about 96 events with rather larger uncertainty of order 35-40%. While one could argue that this is not large discrepancy it does not give good agreement. But there is actually way to test whether we deal almost exclusively with $B^0 \to J/\psi\rho^0$ decay or whether we need to look for something else. If it is exclusive $B^0 \to J/\psi\rho^0$ then this bump should be visible at low dipion masses, but not at higher dipion masses. In Fig. 21 we show invariant mass distributions for events with dipion mass below 980 MeV/c^2 and above. While statistics gets smaller when we split our dataset it is rather clear that bump is not exclusively at small dipion masses, but that there is some other contribution besides $B^0 \to J/\psi\rho^0$. From PDG it seems that $B^0 \to J/\psi\pi^+\pi^-$ branching fraction is about twice as large as $B^0 \to J/\psi\rho^0$, but it does not provide any other observed decay.

Before concluding this part, we can ask ourself whether we deal with real b-hadron or not. If it is b-hadron, then it should have long lifetime. In Fig. 22 we show the invariant mass distribution when in addition to neural network selection we require that L_{xy} significance of dipion vertex is larger than 10. For this plot we do not perform trigger paths selection. It is obvious from this distribution that structure we observe in our distribution is real B^0 signal, but it is hard at this moment to attribute it fully to $B^0 \to J/\psi\rho^0$ decay.

Given all this information we can conclude that structure we observe is correctly reconstructed B^0. As such we should use real b-hadron like signal to describe it instead of adhoc broad Gaussian. From this reason we removed broad Gaussian from background description and use same double Gaussian shape as for $B_s \to J/\psi f_0(980)$ with position of the signal fixed to 5.279 GeV/c^2. As we cannot easily fully attribute it to known decays we leave its normalization free in the fit. Finally as from simulated events we expect to have another rise below 5.26 GeV/c^2, we moved lower fit boundary to 5.26 GeV/c^2 to avoid additional

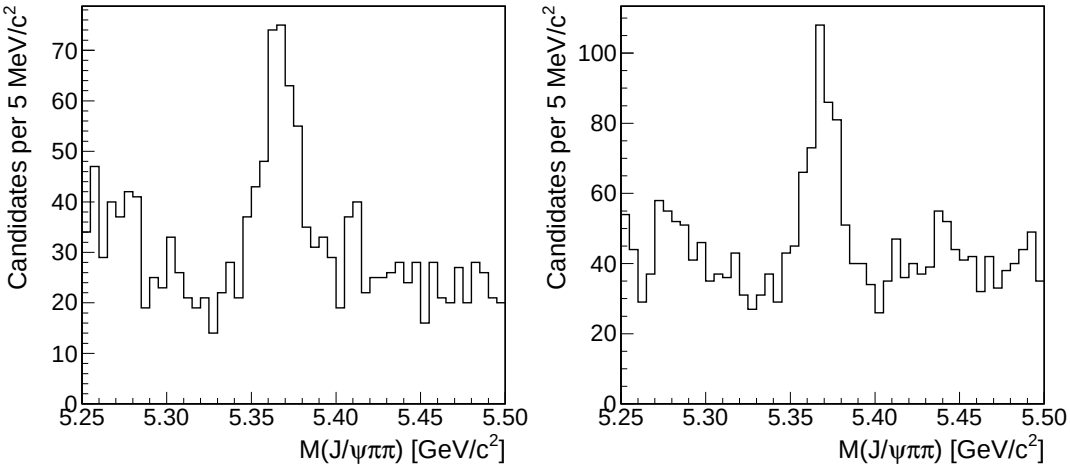

Figure 21: The invariant mass distribution of $J/\psi\pi^+\pi^-$ candidates for events where dipion mass is smaller than 980 MeV/c^2 (left) and larger than 980 MeV/c^2.

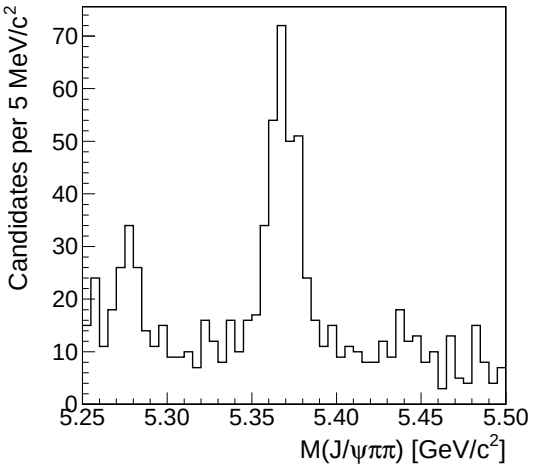

Figure 22: The invariant mass distribution of $J/\psi\pi^+\pi^-$ candidates with additional cut on dipion L_{xy} significance. For this plot we do not perform trigger paths selection.

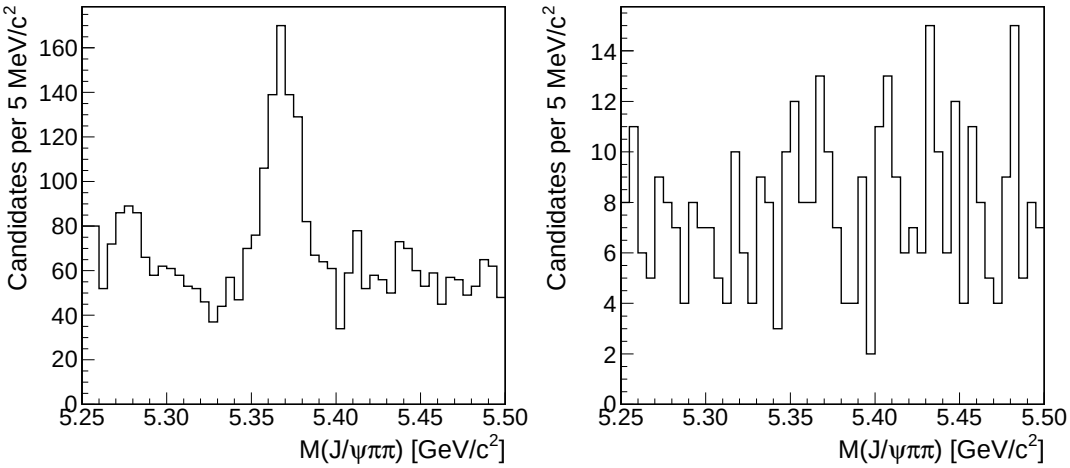

Figure 23: The $J/\psi\pi\pi$ invariant mass distribution for events where only one candidate is selected (left) and events with multiple candidates (right).

complication.

C Multiple Candidates

In Fig. 23 we show $J/\psi\pi\pi$ invariant mass distribution separately for events with single candidate selected and with more than one candidate in event. Corresponding dipion mass distributions are shown in Fig. 24. We conclude that candidates from events with multiple candidates do not show any alarming feature which would require special treatment.

References

[1] Abulencia, A. *et al.* Observation of B_s^0-$\bar{B}_s^0$ oscillations. *Phys. Rev. Lett.* **97**, 242003 (2006). `hep-ex/0609040`.

[2] Aaltonen, T. *et al.* First Flavor-Tagged Determination of Bounds on Mixing-Induced CP Violation in $B_s^0 \to J/\psi\phi$ Decays. *Phys.Rev.Lett.* **100**, 161802 (2008). `arXiv:0712.2397`.

[3] Abazov, V. M. & Others. Measurement of b_s^0 mixing parameters from the flavor-tagged decay $b_s^0 \to j/\psi\phi$. *Phys. Rev. Lett.* **101**, 241801 (2008). `0802.2255`.

[4] Altonen, T. *et al.* An updated measurement of the cp violating phase $\beta_s^{J/\psi\phi}$. *CDF Public Note 9458* (2008).

[5] Stone, S. & Zhang, L. S-waves and the measurement of cp violating phases in bs decays. *Phys. Rev. D* 074024 (2009). `0812.2832`.

[6] Stone, S. S-waves and the extraction of $beta_s$ (2010). * Temporary entry *, `arXiv:1009.4939`.

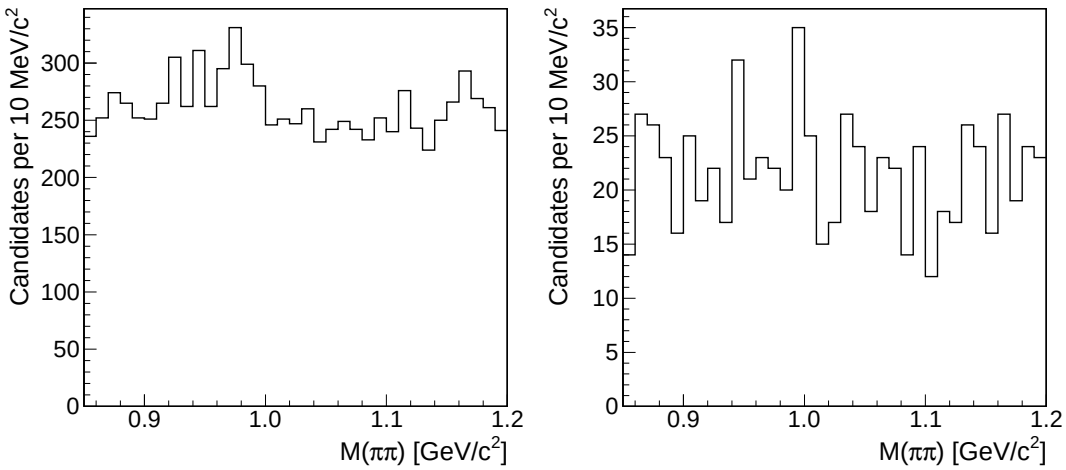

Figure 24: The dipion invariant mass distribution for events where only one candidate is selected (left) and events with multiple candidates (right).

[7] Stone, S. & Zhang, L. Measuring the cp violating phase in b_s mixing using $b_s \to j\psi f_0(980)$ (2009). `0909.5442`.

[8] Yi, K., Nachtman, J., Chung, K., Culbertson, R. & Miao, T. Search for exotic states in the $j/\psi\phi$ mass spectrum in exclusive b decays. *CDF Note 9346* (2008).

[9] F. Azfar, J. Boudreau, M. Feindt, J.P. Fernandez, K. Gibson, G. Giurgiu, G. Gomez-Ceballos, T. Kuhr, M. Kreps, L. Oakes, J. Morlock, M. Paulini, E. Pueschel, A. Schmidt, "An Update of the Measurement of the CP-Violating Phase β_s using $B_s \to J/\psi\phi$ Decays", CDF note 10053.

[10] Nakamura, K. *et al.* Review of particle physics. *J. Phys.* **G37**, 075021 (2010).

[11] Feindt, M. A neural bayesian estimator for conditional probability densities `physics/0402093`.

[12] Feindt, M. & Kerzel, U. The NeuroBayes neural network package. *Nucl.Instrum.Meth.* **A559**, 190–194 (2006).

[13] Abe, F. *et al.* Ratios of bottom meson branching fractions involving J/ψ mesons and determination of b quark fragmentation fractions. *Phys.Rev.* **D54**, 6596–6609 (1996). `hep-ex/9607003`.

[14] Flatte, S. M. Coupled - channel analysis of the pi eta and k anti-k systems near k anti-k threshold. *Phys. Lett.* **B63**, 224 (1976).

[15] Ablikim, M. *et al.* Resonances in $J/\psi \to \phi\pi^+\pi^-$ and ϕK^+K^-. *Phys. Lett.* **B607**, 243–253 (2005). `hep-ex/0411001`.