Search for a Light Scalar in Single Photon Decays of Υ Resonances

Yury G. Kolomensky

Lawrence Berkeley National Laboratory and
University of California, Berkeley

Abstract

We search for a light scalar particle produced in single-photon decays of the Υ resonances through the process $\Upsilon \to \gamma + A^0$, $A^0 \to$ invisible. Such an object appears in supersymmetric extensions of the Standard Model, such as NMSSM, where a light CP-odd Higgs boson naturally couples strongly to b-quarks. If in addition there exist a light stable neutralino, decays of A^0 could be preferentially to an invisible final state. We search for events with a single high-energy photon and a large missing mass, consistent with a 2-body decay of Υ.

Contents

1 Introduction

This document describes an analysis of Run7 data, collected at the $\Upsilon(3S)$ and $\Upsilon(2S)$ resonances. The physics case for the running at resonances below $\Upsilon(4S)$ has been documented in BAD1928 [1], and much of the theoretical motivation is shamelessly borrowed from there.

1.1 Low-mass Higgs boson

The origin of mass is one of the great open questions in physics. The fermion and boson states that form the known matter and interactions in nature are believed to obtain their mass from the spontaneous breaking of an internal symmetry of the Standard Model. This produces mass but also a new state: the Higgs boson. In the simplest scenario, the Standard Model electro-weak gauge symmetry is broken and produces a single observable Higgs boson, which indirect constraints and direct searches suggest should be heavy ($m_H > 114.4\,\text{GeV}/c^2$ from direct searches [2] and $m_H = 129^{+74}_{-49}\,\text{GeV}/c^2$ from electro-weak constraints [3]).

The Standard Model and the simplest electroweak symmetry breaking scenario produce new problems which are to be addressed at the next generation of collider experiments. The mass parameter of the Higgs potential (μ) is not naturally of order the weak scale, and radiative corrections to it are quadratically divergent. Supersymmetry is one model that produces new couplings for the Higgs and regulates the divergence. However, the simplest incarnation of supersymmetry - the Minimal Supersymmetric Standard Model (MSSM) - presents new theoretical problems.

Some theoretical efforts to solve unattractive features of MSSM typically result in models that allow one of the Higgs fields to be light. Its mass can be so low that it is accessible to B factories, yet escaping constraints from LEP and SLC.

1.2 Low-mass Higgs Scenarios

The MSSM solves the problem of quadratic divergence of radiative corrections to the Higgs potential's mass parameter, μ. However, the MSSM fails to explain other features of μ. For instance, there is no motivation as to why the mass scale of the Higgs, μ, should be so small compared to the next natural scale, the Planck scale. Efforts to solve this so-called "μ-problem" have motivated theorists to add one additional singlet Higgs field to the doublet structure of the MSSM - the Next-to-Minimal Supersymmetric Standard Model, or NMSSM [4].

The addition of a singlet Higgs field to solve the μ problem leads to a testable consequence: there is a CP-odd Higgs boson with a largely singlet composition (as the Higgses can mix) whose mass is not required to be large. Such a state would have evaded detection at LEP and SLD for several reasons. In models with many degrees of freedom, such as the NMSSM, the coupling to the Z^0 is not required to be large, and so a low-mass Higgs state could evade the constraints from LEP and SLD on the existence of new particles with masses less than M_{Z^0}. In addition, the Standard Model Higgs may have avoided detection as its dominant decay mode would be $H^0 \to A^0 A^0$, where A^0 is the CP-odd low-mass Higgs. If the A^0 were light enough, such that $A^0 \to b\bar{b}$ was not possible, then the dominant visible final state of SM Higgs decay would be 4τ. There are no constraints from LEP on this final state.

Such a Higgs particle should be detectable at the B-factories. The coupling to the b-quark will still be significant and an ideal place to search for the Higgs would be $\Upsilon \to \gamma A^0$, as originally

proposed by Wilczek [5]. A study of the NMSSM parameter space [6] has produced predictions for the branching fraction to this final state (Fig. 1).
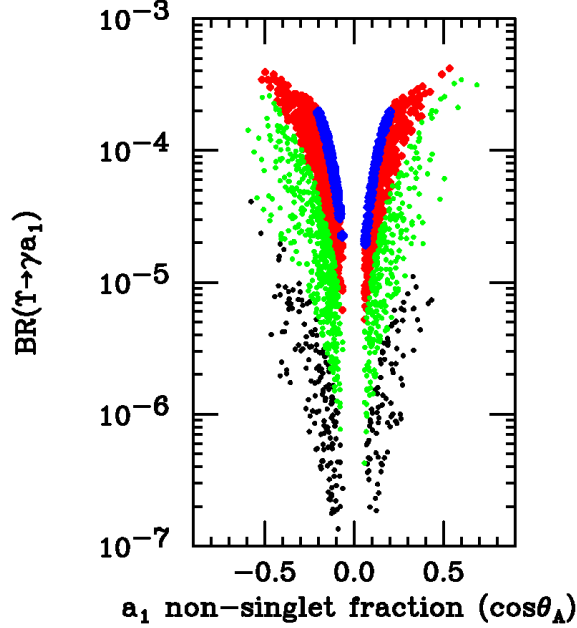

Figure 1: Parameter space scan of the NMSSM yielding predictions for $\mathcal{B}(\Upsilon \to \gamma A^0)$ for many scenarios. Points shaded dark (blue) are $m_{A^0} < 2m_\tau$, those shaded medium gray (red) satisfy $2m_\tau < m_{A^0} < 7.4\,\mathrm{GeV}/c^2$, those shaded light gray (green) satisfy $7.5\,\mathrm{GeV}/c^2 < m_{A^0} < 8.8\,\mathrm{GeV}/c^2$, and those shaded black satisfy $8.8\,\mathrm{GeV}/c^2 < m_{A^0} < 9.2\,\mathrm{GeV}/c^2$. This figure is taken from [6]. The branching fraction depends also on the fraction of the A^0 which is non-singlet (abscissa). On the graph, the lightest CP-odd Higgs state A^0 is labeled a_1 (not to be confused with a known unflavored meson).

1.2.1 Kinematics and Decays of a Low-Mass Higgs

The cleanest signature of the $\Upsilon \to \gamma A^0$ decay is a monochromatic photon with energy

$$E_\gamma^* = \frac{m_\Upsilon^2 - m_{A^0}^2}{2m_\Upsilon} \tag{1}$$

recoiling against a low-multiplicity final state. Running on $\Upsilon(3S)$ resonance, we can look for the radiative $\Upsilon(3S)$ decays directly, or for the decays of a lower-mass $\Upsilon(1S)$ or $\Upsilon(2S)$, produced through the di-pion transition from $\Upsilon(3S)$. Likewise, at the $\Upsilon(2S)$ resonance, both direct radiative decays and the radiative decays of $\Upsilon(1S)$ (tagged by the di-pion transition from $\Upsilon(2S)$) are possible. Due to the sizable branching fraction for $\Upsilon(2S) \to \pi\pi\Upsilon(1S)$ and an additional kinematic constraint from the di-pion system, the radiative $\Upsilon(1S)$ decays offer a competitive data sample.

The decays of the light Higgs boson depend on its mass and couplings. The decays into tau pairs is often considered a "golden" channel for its relatively clean final state (a pair of acollinear

leptons, *e.g.*, $e\mu$) and identifiable kinematic signatures (missing energy and momentum). Such decays would be allowed if $m_{A^0} > 2m_\tau$.

If the mass of A^0 is above the threshold for the hadronic transitions, $m_{A^0} > 3m_\pi$, the hadronic final states are possible. If $m_{A^0} < 2m_\tau$, the dominant Standard Model leptonic decay mode would be to muon pairs. According to the analysis of Ref. [7], though somewhat problematic theoretically, such a scenario may explain an observed excess of positrons in the galactic center and a hint of a narrow state decaying to $\mu^+\mu^-$ observed by the HyperCP experiment [8], if $212\,\mathrm{MeV} < m_{A^0} < 420\,\mathrm{MeV}$.

1.3 Invisible Decays of A^0

In certain NMSSM scenarios, particularly those in which the mass of the lightest supersymmetric particle is above m_τ or if $m_{A^0} < 2m_\tau$, the dominant decay mode of A^0 may be invisible: $A^0 \rightarrow \chi\bar{\chi}$. The cleanest experimental signature of such decays is production of nearly monochromatic single photons in decays $\Upsilon \rightarrow \gamma A^0$, accompanied by a significant missing energy/momentum.

The current best limit on the branching fraction $\mathcal{B}(\Upsilon \rightarrow \gamma X)$ with $X \rightarrow$ invisible comes from a CLEO measurement on $\Upsilon(1S)$ in 1994 [9]. The quoted limits range from 1.3×10^{-5} for the lightest m_X (highest-energy photons) to $(4\text{–}6) \times 10^{-4}$ for $m_X \approx 8\,\mathrm{GeV}/c^2$ (PDG quotes this result as $\mathcal{B}(\Upsilon(1S) \rightarrow \gamma X) < 3 \times 10^{-5}$ for $m_X < 7.2\,\mathrm{GeV}/c^2$ [10]). There are currently no competitive measurements at the higher-mass Υ resonances.

The events $\Upsilon \rightarrow \gamma A^0$, $A^0 \rightarrow$ invisible are characterized by a single photon with an energy given by Eq. (1). Detection of such events in *BABAR* requires special single-photon triggers and filters, which were installed in September 2007, and modified during Run7.

In the following sections, we describe the search for the radiative transitions $\Upsilon \rightarrow \gamma A^0$, followed by invisible decays $A^0 \rightarrow$ invisible. We will focus on the direct $\Upsilon(3S)$ and $\Upsilon(2S)$ decays, leaving the di-pion transitions for $\Upsilon(1S)$ for a subsequent analysis.

2 Trigger Lines

Detection of the low-multiplicity single photon events requires dedicated trigger and filter lines. Event processing and selection proceeds in three steps: in L1, L3, and in RecoBGFilter.

2.1 Level-1 Trigger

The line responsible for passing the majority of the events of interest is "`1E`", which requires an single L1 cluster with the energy above 800 MeV in the Lab. This line was activated in L1 on 26-AUG-2007, before run **76135**. Prior to that, only a short L1PASSTHRU run (57920) was taken with 1E line activated.

2.2 Level-3 Trigger

Three lines have been developed by the trigger group to pass the events of interest:

- `L3OutSingleGamma`: requires a single isolated EMC cluster and no L3 tracks consistent with originating from the IP. The EMC cluster threshold is $E_\gamma^* > 2$ GeV. This line was activated before run **76135**.

- `L3OutSingleGammaLowE`: requires a single isolated EMC cluster and no L3 tracks consistent with originating from the IP. The EMC cluster threshold is $E_\gamma^* > 1$ GeV. This line was activated before run **78308**.

- `L3OutSingleGammaVeryLowE`: requires a single isolated EMC cluster and no L3 tracks consistent with originating from the IP. The EMC cluster threshold is $E_\gamma^* > 0.6$ GeV. This line was activated before run **78308**. This line was primarily conceived to allow soft-photon transitions $\Upsilon(3S) \to \gamma\gamma\Upsilon(1S)$, followed by invisible decays of $\Upsilon(1S)$. Since it also requires `1E` L1 line, it is only about 50% efficient. The rate is dominated by QED backgrounds.

The requirements for the "isolated cluster" requires that there are no tracks within $|\Delta\phi| < 10°$ and $\Delta\theta < 5°$ of the cluster at the interception with the EMC. The IP tracks are defined by $|d_0| < 1.5$ cm, $|z_0| < 10$ cm, and $p_T > 250$ MeV.

2.3 DigiFilter

L3 lines above are copied as DigiFilter lines `DigiFSingleGamma`, `DigiFSingleGammaLowE`, and `DigiFSingleGammaveryLowE`, with unit prescale. Lines `DigiFSingleGammaLowE` and `DigiFSingleGammaveryLowE` were activated before run 78308.

2.4 RecoBGFilter

Two RecoBGFilter lines enable the processing of the data in Elf:

- `BGFSingleGammaInvisible`: requires at most one photon (EMC cluster) above $E_\gamma^* > 3$ GeV, no tracks above $p^* > 1$ GeV. Activated before run **76135**. The EMC energy scale for the clusters is biased in RecoBGFilter, so the energy cut corresponds to the requirement on the EMC bump energy of 3.2 GeV.

- `BGFSingleGammaInvisibleLowE`: requires at most one photon (EMC cluster) above $E_\gamma^* > 1.5$ GeV, no tracks above $p^* > 100$ MeV. The cluster is required to pass loose shower quality cuts: $5 < N_{\mathrm{crys}} < 61$, $LAT < 0.8$, $a_{42} < 0.15$. These cuts are looser than the corresponding offline selections. `BGFSingleGammaInvisibleLowE` was activated before run **76135**. In terms of the EmcBump energy, this selection corresponds to $E_\gamma^* > 1.8$ GeV.

The selector module is **RecoSingleGammaSelector**.

3 Data and Monte Carlo Samples

We use the following data samples in the analysis[1]:

- Run6 data taken with the single photon trigger activated, *i.e.* data sample `AllEvents-Run6-OnPeak-R22`, starting with run 76135. This sample amounts to 3864 pb^{-1} taken at $\Upsilon(4S)$.

---

[1]We use `BbkLumiTools V00-00-36`, which removes the factor of 1.068 from the "L3" estimate of the luminosity for Run7

- Run7 `AllEvents` data collected on $\Upsilon(3S)$ resonance: data samples
  `AllEvents-Run7-R24-Y3S-OnPeak-Low` ($1200 \, \text{pb}^{-1}$) and
  `AllEvents-Run7-R24-Y3S-OnPeak-High` ($1282 \, \text{pb}^{-1}$).

- Run7 `AllEvents` data collected off $\Upsilon(3S)$ resonance: data sample
  `AllEvents-Run7-R24-Y3S-OffPeak-v02` ($2415 \, \text{pb}^{-1}$).

- Run7 `BGFilterSkim` data collected on $\Upsilon(3S)$ resonance: data sample
  `BGFilterSkim-Run7-Y3S_OnPeak-R24-v01` ($19,543 \, \text{pb}^{-1}$ collected after run 78307).
  For preliminary tests, we use a sample of $2393 \, \text{pb}^{-1}$ in the run range 79441-79949, which
  overlaps with the
  `AllEvents-Run7-R24-Y3S-OnPeak-High` sample.

- Run7 `BGFilterSkim` data collected off $\Upsilon(3S)$ resonance: data sample
  `BGFilterSkim-Run7-Y3S_OffPeak-R24-v01` ($2648 \, \text{pb}^{-1}$).

- Run7 `AllEvents` data collected on $\Upsilon(2S)$ resonance: data sample
  `AllEvents-Run7-Y2S_OnPeak-R24`. For preliminary tests, we use a sample of $1297 \, \text{pb}^{-1}$
  in the run range 80826-80929.

- Run7 `AllEvents` data collected off $\Upsilon(2S)$ resonance: data sample
  `AllEvents-Run7-Y2S_OffPeak-R24` ($973 \, \text{pb}^{-1}$).

- Run7 `AllEvents` data collected on $\Upsilon(4S)$ resonance in Dec. 2007: sample
  `AllEvents-Run7-OnPeak-R24` ($881 \, \text{pb}^{-1}$).

- Run7 `FinalScan` data collected above $\Upsilon(4S)$ resonance at the end of Run 7: sample
  `AllEvents-Run7-FinalScan-v01` ($4544 \, \text{pb}^{-1}$).

We also use the following SP-10 Monte Carlo modes for signal and background characterization:

- SP-8757 ($\Upsilon(3S) \to \gamma A^0 (0.1 \text{ GeV})$): 162k events (113k after Jan2008b)

- SP-8756 ($\Upsilon(3S) \to \gamma A^0 (1 \text{ GeV})$): 162k events (113k after Jan2008b)

- SP-8740 ($\Upsilon(3S) \to \gamma A^0 (2 \text{ GeV})$): 162k events (113k after Jan2008b)

- SP-8919 ($\Upsilon(3S) \to \gamma A^0 (3 \text{ GeV})$): 153k events (107k after Jan2008b)

- SP-8741 ($\Upsilon(3S) \to \gamma A^0 (4 \text{ GeV})$): 162k events (113k after Jan2008b)

- SP-8920 ($\Upsilon(3S) \to \gamma A^0 (5 \text{ GeV})$): 153k events (107k after Jan2008b)

- SP-8742 ($\Upsilon(3S) \to \gamma A^0 (6 \text{ GeV})$): 162k events (113k after Jan2008b)

- SP-8921 ($\Upsilon(3S) \to \gamma A^0 (7 \text{ GeV})$): 153k events (107k after Jan2008b)

- SP-8743 ($\Upsilon(3S) \to \gamma A^0 (8 \text{ GeV})$): 162k events (113k after Jan2008b)

- SP-1074 ($e^+ e^- \to \gamma \gamma$ at $\sqrt{s} = 10.355 \text{ GeV}$): sample
  `SP-1074-Run7-Y3S_OnPeak-R24-v01`, 31,808k events (equivalent to $13.8 \, \text{fb}^{-1}$).

- SP-3981 ($e^+ e^- \to \mu^+ \mu^- (\gamma)$ at $\sqrt{s} = 10.355 \text{ GeV}$): 17,040k events (equivalent to $17.4 \, \text{fb}^{-1}$).

- SP-3429 ($e^+ e^- \to \tau^+ \tau^- (\gamma)$ at $\sqrt{s} = 10.355 \text{ GeV}$): 3,016k events (equivalent to $6.2 \, \text{fb}^{-1}$).
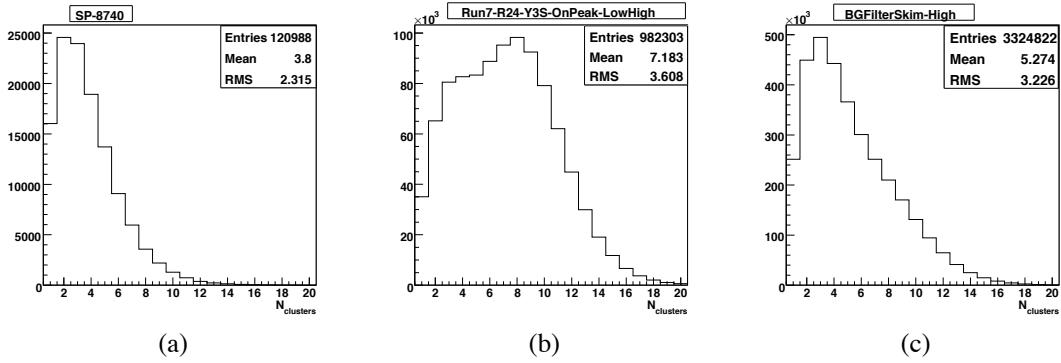
Figure 2: EMC cluster multiplicity for the pre-selected events in (a): signal Monte Carlo sample ($m_{A^0} = 2\,\text{GeV}$) `SP-8740`; (b): data ''`LowHigh`'' sample ($2.5\,\text{fb}^{-1}$); (c): data ''`BGFilterSkim-High`'' sample ($2.4\,\text{fb}^{-1}$).

4 Event Selection

4.1 Combinatorics and pre-selection

We convert events into ntuple form with the `BtaTupleMaker V00-03-23` package. The first step is the custom module `BetaTools/BtaMakeOneCand`, which combines all tracks (from `Outer` list) and neutral EMC clusters (`CalorNeutral` list) in the event into a single Υ candidate. `BtaTupleMaker` is then used to convert the candidate and its daughters into ntuple form. We store up to 20 EmcCands per event. We also store the `NeutralHad` list, which contains the neutral IFR clusters. Optionally, we can also store the data on the un-associated hits from the DIRC and the DCH (this requires the "refit" mode, and is not currently done, although we envision staging the hit data for a few select candidate events if they require additional inspection).

The following tag-level cuts are implemented at the processing stage:

- `nTracks` $= 0$

- $0 <$ `nNeutralCands` < 21

- $1.5 <$ `e1Mag` < 6.0

- `e2Mag` < 1.0

The last requirement makes a cut on the energy of the 2nd highest (in CMS) energy neutral cluster in the event. The cluster multiplicity for various event samples (data and signal Monte Carlo) is shown in Fig. 2.

4.2 Selection Variables

A limited number of variables is available for our very low-multiplicity event samples. We use the following selection criteria to select the events of interest:

- RecoBGFilter bits: `BGFSingleGammaInvisible` and `BGFSingleGammaInvisible` lines.

- Photon quality: number of crystals in the EmcBump N_{crys} and transverse moments LAT and a_{42}

- Fiducial selection of primary photons: $\cos\theta_\gamma$ and ϕ_γ. This tends to select photons in the barrel part of the EMC.

- Extra particles in the event: require $N_{\mathrm{tracks}} = 0$ (counting `Outer` tracks, which includes photon conversions). We also apply cuts on the energy of the second-highest photon in the event E_2^* (`e2Mag` in the tag database, computed in CMS), extra energy in the calorimeter E_{extra}, computed in the Lab and the ϕ angle difference between the primary and the second photon in the event $\phi_2 - \phi_1$. Leaving some non-zero calorimeter energy allows for the presence of machine backgrounds. The cut on $\phi_2 - \phi_1$ suppresses $\gamma\gamma$ and other QED background (*e.g.*, FSR radiation from bhabhas).

- IFR veto: we cut on the ϕ angle difference between the primary photon and the nearest IFR cluster (from `NeutralHad` list). This variable, $\Delta\phi_{\mathrm{NH}}$, rejects the $e^+e^- \to \gamma\gamma$ events in which one of the photons is lost in the dead regions between the EMC crystals.

4.3 Selection Optimization

We optimize the selection in two broad energy ranges: $3.2 < E_\gamma^* < 5.5$ GeV, which corresponds to the `BGFSingleGammaInvisible` line. This region is dominated by the $e^+e^- \to \gamma\gamma$ background, especially near $E_\gamma^* = \sqrt{s}/2$, where $\gamma\gamma$ events peak. The selection is optimized to reduce this peaking background as much as possible.

The second energy range is $1.8 < E_\gamma^* < 3.7$ GeV, which corresponds to the `BGFSingleGammaInvisibleLowE` line. This region is dominated by the low-angle radiative bhabha events $e^+e^- \to e^+e^-\gamma$, in which both electron and positron disappear in the beampipe. In the region $3.0 < E_\gamma^* < 3.7$ GeV, the tail from the $e^+e^- \to \gamma\gamma$ background is significant.

We optimize the selection iteratively, maximizing $S/\sqrt{B}$. The signal efficiency S is computed from at least two signal Monte Carlo samples with the photon energy in the appropriate range. The background B is taken from the $\Upsilon(3S)$ "High" sample. Each cut is first optimized independently on the pre-selection sample, and then varied around the chosen value with all the other cuts set to their optimal values.

An example of such optimization for two variables, E_2^* and N_{crys}, is shown in Fig. 3-4.

4.4 IFR Veto

The largest background in the high-energy region, $3.2 < E_\gamma^* < 5.5$ GeV, comes from the QED process $e^+e^- \to \gamma\gamma$, in which one of the photons is lost. Even after the fiducial selection that requires that the missing photon point to the EMC barrel, 1% of the $e^+e^- \to \gamma\gamma$ events contain only one reconstructed EMC cluster. Most often, this happens when the second photon hits a dead region between the EMC crystals (crystals are oriented to be projective in ϕ). This is evident in Fig. 5. Fortunately, in such events the photon is likely to deposit significant energy
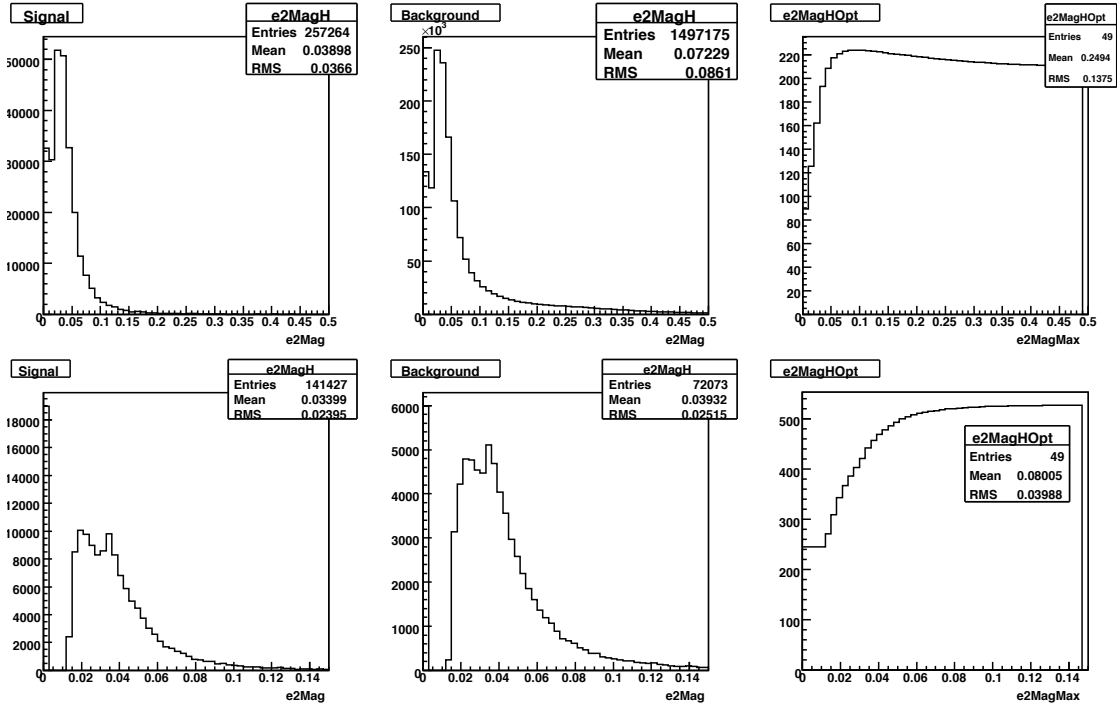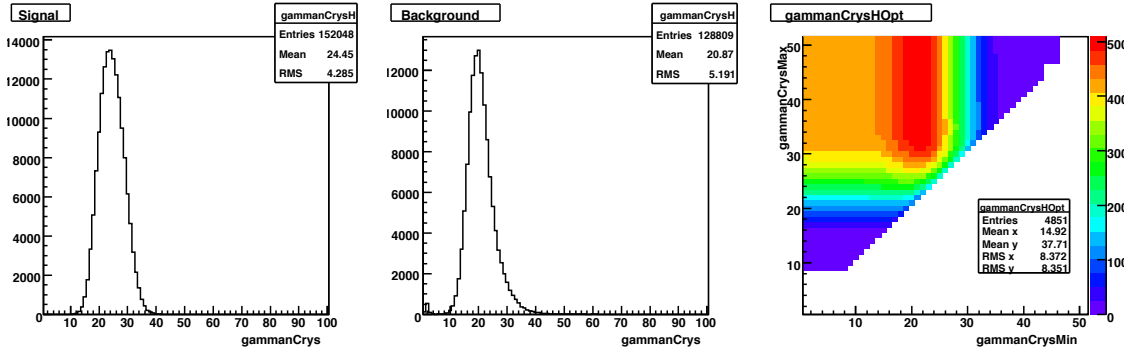
Figure 3: Optimization of the cut for the E_2^* variable in the primary photon energy range $1.8 < E_\gamma^* < 3.7$ GeV. The optimization searches for the value $E_{2\,\mathrm{max}}^*$ for the selection that requires $E_2^* < E_{2\,\mathrm{max}}^*$. Top: loose pre-selection, bottom: tight selection where all variables except E_2^* are near their optimal values. The left panel shows the distribution for the signal events, the middle panel shows the distribution for the data, and the right panel shows $S/\sqrt{B}$ as a function of $E_{2\,\mathrm{max}}^*$.

Figure 4: Optimization of the cut for the N_{crys} variable in the primary photon energy range $1.8 < E_\gamma^* < 3.7$ GeV. The optimization searches for a pair of values $[N_{\mathrm{crys\,min}}, N_{\mathrm{crys\,max}}]$ for the selection that requires $N_{\mathrm{crys\,min}} \leq N_{\mathrm{crys}} \leq N_{\mathrm{crys\,max}}$. Tight selection, where all variables except N_{crys} are near their optimal values, is applied. The left panel shows the distribution for the signal events, the middle panel shows the distribution for the data, and the right panel shows $S/\sqrt{B}$ as a function of the cut values.

11
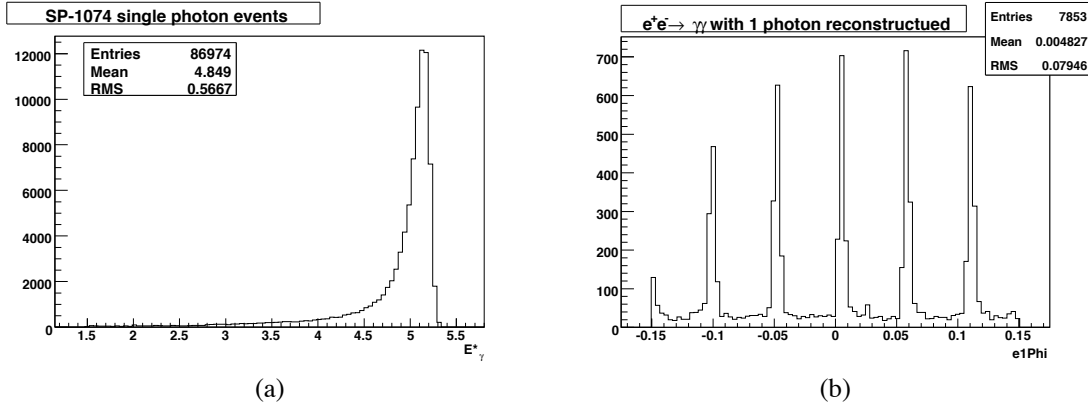
Figure 5: a): Distribution of E^*_γ for $e^+e^- \to \gamma\gamma$ Monte Carlo events where only a single photon is observed (the events pass the standard selection cuts except for the IFR veto). The number of events corresponds to the EMC detection inefficiency for the second photon of about 1%. (b): Distribution of the ϕ_1 angle of the primary photon (region $|\phi_1| < 0.15$ is shown). The peaks correspond to the projective boundaries between the EMC crystals.

in the IFR, resulting in a cluster. We therefore veto such events by requiring that there are no `NeutralHad` candidates in the IFR back-to-back in ϕ with the primary photon.

Fig. 6 shows the distribution of $\cos(\Delta\phi_{\mathrm{NH}}) = \cos(\phi_\gamma - \phi_{\mathrm{NH}})$, the difference in the ϕ angle between the primary photon and the farthest (in ϕ) neutral hadron. The correlation in the direction of the primary photon and the neutral hadron (in the location where the second photon should have been) is apparent as a large peak at $\cos(\Delta\phi_{\mathrm{NH}}) = -1$. There is also a correlation in the forward direction ($\cos(\Delta\phi_{\mathrm{NH}}) = 1$) due to the shower leakage. A cut $\cos(\Delta\phi_{\mathrm{NH}}) > -0.9$ rejects 97% of the $e^+e^- \to \gamma\gamma$ candidates with a single EMC cluster[2]

The only remaining issue is the gap between the IFR sectors, visible in the distribution of the ϕ_γ (Fig. 7). The six spikes in the rate of events passing the IFR veto correspond to the dead regions between the six IFR sectors, where the veto is inefficient. To combat that, we apply an additional fiducial requirement, $\cos(6\phi_\gamma) < 0.96$.

4.5 Final Selection

The final selection for the energy range $3.2 < E^*_\gamma < 5.5$ GeV is summarized in Table 1. A plot of the figure of merit, $S/\sqrt{B}$ for each successive cut, is shown in Fig. 8. Selection efficiency as a function of the mass of A^0, determined from signal Monte Carlo, is shown in Fig. 9. The selection efficiency for signal is about 11%. Most of the efficiency loss occurs from the fiducial requirements and the IFR veto.

The "N-1" plots of the selection variables, with all cuts except on the variable being plotted applied, are shown in Appendix A (Fig. 65-66).

The final selection for the energy range $1.8 < E^*_\gamma < 3.7$ GeV is summarized in Table 2. Since this region is dominated by the radiative bhabha events (which have no peaking structure),

---

[2]Technically, this requirement only works in the IFR barrel with relatively precise ϕ measurements. However, this matches our EMC fiducial requirements very well. We currently do not use the endcaps due to excessive backgrounds from QED processes.
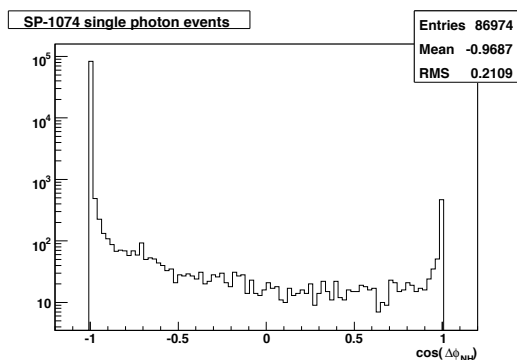
Figure 6: Distribution of $\cos(\Delta\phi_{\mathrm{NH}})$ for $e^+e^- \to \gamma\gamma$ Monte Carlo events passing with a single EMC cluster.

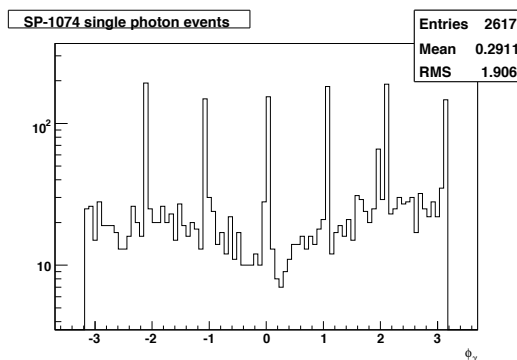Figure 7: Distribution of ϕ_γ for $e^+e^- \to \gamma\gamma$ Monte Carlo events passing the IFR veto. The six regions with highly increased rate correspond to the dead regions between the IFR sectors in the barrel.

Table 1: Selection criteria for the high energy region, $3.2 < E_\gamma^* < 5.5$ GeV.

Variable	Cut
BGFilter	`BGFSingleGammaInvisible`
Number of crystals in EmcCand	$20 < N_{\mathrm{crys}} < 48$
LAT shower shape	$0.24 < LAT < 0.51$
a_{42} shower shape	$a_{42} < 0.07$
Polar angle acceptance	$-0.31 < \cos\theta_\gamma^* < 0.6$
2nd highest cluster energy (CMS)	$E_2^* < 0.2$ GeV
Extra photon correlation	$\cos(\phi_2 - \phi_1) > -0.95$
Extra EMC energy (Lab)	$E_{\mathrm{extra}} < 0.1$ GeV
IFR veto	$\cos(\Delta\phi_{\mathrm{NH}}) > -0.9$
IFR fiducial	$\cos(6\phi_\gamma) < 0.96$

13
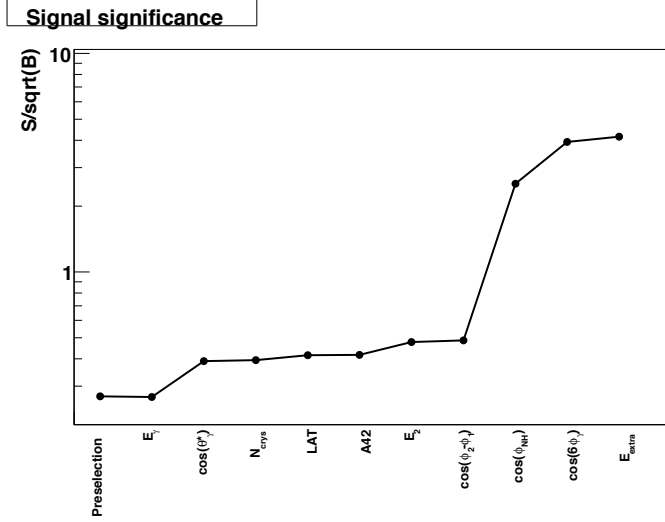
Figure 8: Optimization figure of merit, $S/sqrtB$, in the high energy region $3.2 < E_\gamma^* < 5.5$ GeV. We assume the full $\Upsilon(3S)$ sample ($N_{\Upsilon(3S)} = 120 \times 10^6$), and the signal branching fraction of $\mathcal{B} = 10^{-5}$. Signal efficiency is computed for $m_{A^0} = 2$ GeV (`SP-8740`), and the background efficiency from the ``LowHigh'' sample (2.5 fb^{-1}). The full interval $3.2 < E_\gamma^* < 5.5$ GeV is included (so the actual significance is underestimated).
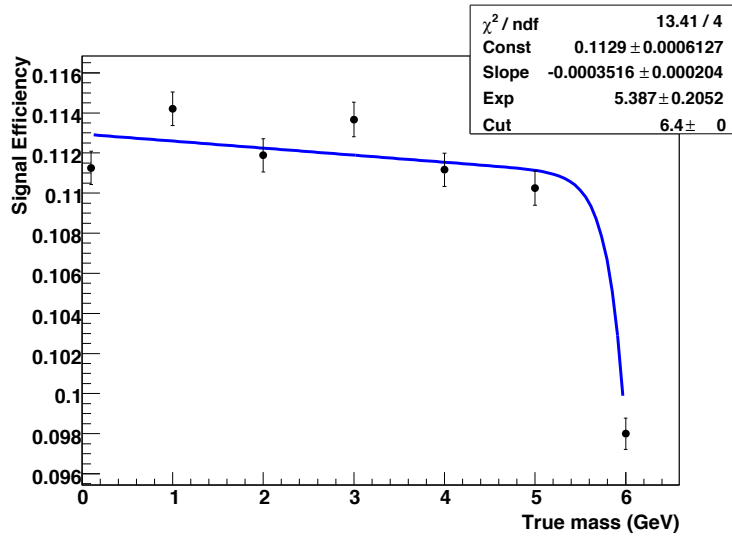
Figure 9: Selection efficiency for the energy range $3.2 < E_\gamma^* < 5.5$ GeV, which corresponds to the A^0 mass range $m_{A^0} < 6.3$ GeV at $\Upsilon(3S)$. The analytical function includes an exponential rolloff of the efficiency due to the energy cut, which takes into account the tail of the γ energy resolution function.

14

Table 2: Selection criteria for the low energy region, $1.8 < E_\gamma^* < 3.7$ GeV.

Variable	Cut
BGFilter	`BGFSingleGammaInvisibleLowE`
Number of crystals in EmcCand	$12 < N_{\mathrm{crys}} < 36$
LAT shower shape	$0.15 < LAT < 0.49$
a_{42} shower shape	$a_{42} < 0.07$
Polar angle acceptance	$-0.6 < \cos\theta_\gamma^* < 0.6$
2nd highest cluster energy (CMS)	$E_2^* < 0.14$ GeV
Extra photon correlation	$\cos(\phi_2 - \phi_1) > -0.95$
Extra EMC energy (Lab)	$E_{\mathrm{extra}} < 0.22$ GeV
IFR veto	$\cos(\Delta\phi_{\mathrm{NH}}) > -0.95$

the optimization prefers high efficiency for the signal. In particular, we apply only a mild IFR veto, which helps reduce the background from the tail of the $e^+e^- \rightarrow \gamma\gamma$ background above $E_\gamma^* > 3$ GeV. The fiducial cuts are also significantly looser than in the high-energy region. The "N-1" plots for the low-energy region are shown in Fig. 69-70.

A plot of the figure of merit, $S/\sqrt{B}$ for each successive cut, is shown in Fig. 10. Selection efficiency as a function of the mass of A^0, determined from signal Monte Carlo, is shown in Fig. 11. The selection efficiency for signal is about 28%. Here we had to be careful with the cut on N_{crys} which is correlated to the energy of the photon in the lab (and therefore, to the photon energy and polar angle in CMS). The selection is looser than the formal optimization would prefer, but produces the efficiency that is nearly independent of energy.

5   Yield Fit in the High Energy Region

The high energy region $3.2 < E_\gamma^* < 5.5$ GeV is dominated by the backgrounds from $e^+e^- \rightarrow \gamma\gamma$ process. The angular distribution for the photons from $e^+e^- \rightarrow \gamma\gamma$ is $1 + \cos^2\theta_\gamma^*$, identical to that of the signal $\Upsilon(3S) \rightarrow \gamma A^0$ decays. Therefore, the polar angle of the photon carries little discriminating power. For the extraction of the signal yield in the data, it is sufficient to perform a one-dimensional fit to the distribution of events in E_γ^*, or, equivalently, to the distribution of the recoil mass m_X^2, computed according to Eq. (1).

Our fitting strategy is as follows. We fit the distribution of m_X^2 to a combination of 3 PDFs:

(a) Signal PDF, determined from Monte Carlo, with floated signal yield. The EMC resolution depends on photon energy, therefore the parameters of the signal PDF depend on the assumed true mass of A^0. We determine the PDFs for a set of generated Monte Carlo masses, and then interpolate the PDFs as a function of assumed mass of A^0.

(b) Peaking background from $e^+e^- \rightarrow \gamma\gamma$. This PDF is determined in data from events that pass the final selection (including fiducial requirements) *except* for the IFR veto cut on $\cos\Delta\phi_{\mathrm{NH}}$. The normalization of this PDF is fixed in the data fit (see below).

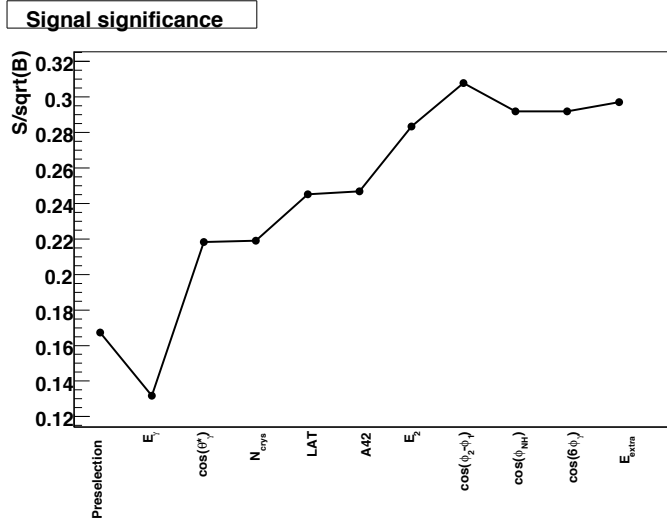(c) "Continuum" background from radiative bhabhas and degraded $e^+e^- \rightarrow \gamma\gamma$ events. This

15

Figure 10: Optimization figure of merit, $S/sqrtB$, in the low energy region $1.8 < E_\gamma^* < 3.7$ GeV. We assume the full `BGFilterSkim` sample ($N_{\Upsilon(3S)} = 82 \times 10^6$), and the signal branching fraction of $\mathcal{B} = 10^{-5}$. Signal efficiency is computed for $m_{A^0} = 7$ GeV (`SP-8921`), and the background efficiency from the ''`BGFilterSkim-High`'' sample (2.4 fb^{-1}). The full interval $1.8 < E_\gamma^* < 3.7$ GeV is included (so the actual significance is underestimated).
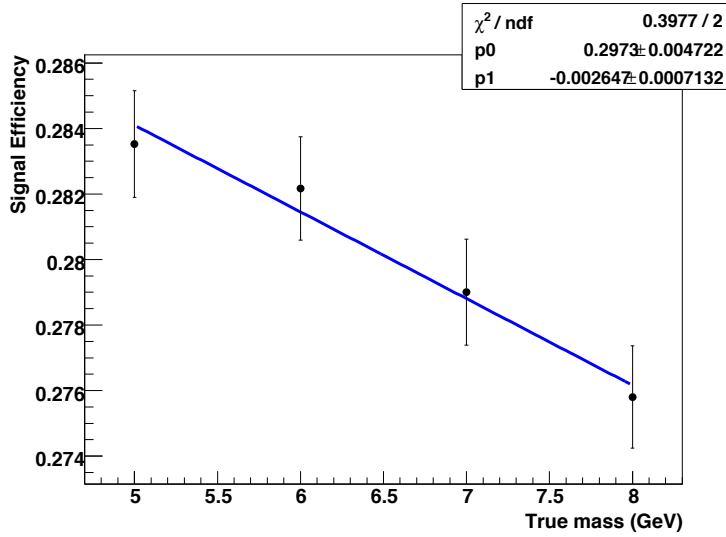
Figure 11: Selection efficiency for the energy range $1.8 < E_\gamma^* < 3.7$ GeV, which corresponds to the A^0 mass range $5.5 < m_{A^0} < 8.7$ GeV at $\Upsilon(3S)$.

16

background is parameterized by a smooth (exponential) function with the shape parameter floated in the data fit.

We fit the data in steps of m_{A^0}, floating the yields of the signal and continuum background events (but fixing the $e^+e^- \to \gamma\gamma$ yield). The final result would be the value of the upper limit on the branching fraction $\mathcal{B}(\Upsilon(3S) \to \gamma A^0)$ as a function of m_{A^0} (or, in a case of a positive observation, the observed mass and the branching ratio).

5.1 Signal PDF

The signal PDF is parameterized as a Crystal Ball function in recoil mass squared m_X^2:

$$
\begin{aligned}
f(m_X^2) &\propto
\begin{cases}
\exp(-t^2/2), \ t \leq -\alpha \\
\left(\frac{n}{|\alpha|}\right)^n \exp(-\alpha^2/2) \left(\frac{n}{|\alpha|} - |\alpha| + t\right)^{-n}, \ t > -\alpha
\end{cases} \\
t &\equiv \frac{m_X^2 - \mu}{\sigma}
\end{aligned}
$$

where $\mu, \sigma, \alpha,$ and n are the PDF parameters. The parameters of the Crystal Ball function depend on the true photon energy, and therefore, on the true value of m_{A^0}. Appendix B shows the fits for the different Monte Carlo samples. We fit the dependence of the Crystal Ball parameters to quadratic functions in m_{A^0} (this corresponds to the dependence of the resolution function on the photon energy E_γ^* and its square root). These fits are shown in Fig. 12.

5.2 Normalization of $e^+e^- \to \gamma\gamma$ Yield

Modeling the main peaking background from $e^+e^- \to \gamma\gamma$ events with a mis-reconstructed or missing photon involves intricate understanding of the EMC geometry, especially the dead space between the crystals. We are aiming for the level of $e^+e^- \to \gamma\gamma$ event rejection of a few parts in 10^6. At this level, it is hard to expect the MC simulations to be reliable. Therefore, we need to be able to estimate the $e^+e^- \to \gamma\gamma$ background from data.

Fortunately, the rejection of the $e^+e^- \to \gamma\gamma$ events factorizes nicely into the EMC-based rejection (based on E_{extra} and E_2^*) and the IFR-based rejections. We can measure the number of events $N_{\gamma\gamma\,EMC}$ passing the EMC-based cuts in the data sample by applying all the cuts (including the fiducial selection) *except* for the IFR veto. This large sample can also be used to determine the PDF of the peaking $e^+e^- \to \gamma\gamma$ background. Then, we measure the rejection power of the IFR veto in a off-resonance control sample

$$
\varepsilon_{\text{IFR}} = \frac{N_{\gamma\gamma}^{\text{off}}}{N_{\gamma\gamma\,\text{EMC}}^{\text{off}}}
\tag{2}
$$

where $N_{\gamma\gamma}^{\text{off}}$ is the number of events passing the final selection (including the IFR veto) in an off-resonance data sample. We use any available off-resonance (that is, any data taken away from $\Upsilon(2S)$ and $\Upsilon(3S)$) sample, and then compute the number of events expected from $e^+e^- \to \gamma\gamma$ background in an on-resonance sample as

$$
N_{\gamma\gamma}^{\text{on}} = \langle \varepsilon_{\text{IFR}} \rangle N_{\gamma\gamma\,\text{EMC}}^{\text{on}}
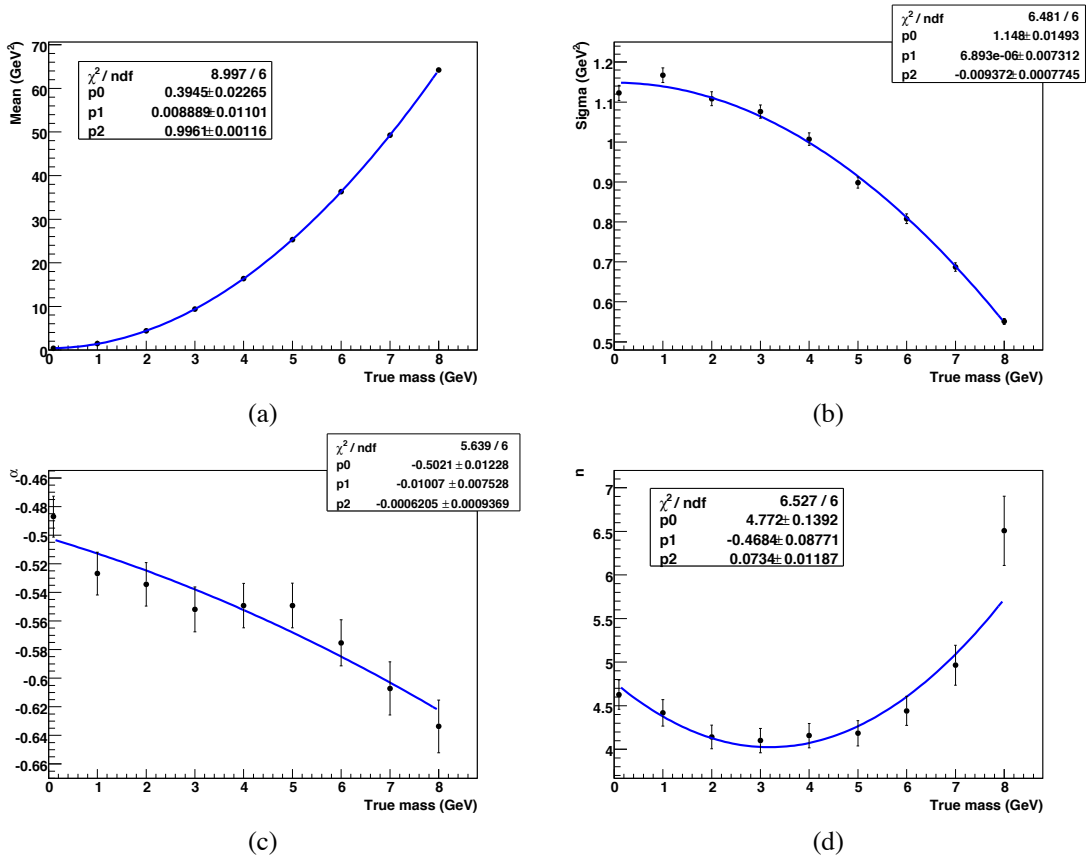\tag{3}
$$

(a)

(b)

(c)

(d)

Figure 12: Dependence of the signal PDF parameters on the true mass m_{A^0}. (a) μ (mean of the Gaussian part of the Crystal Ball function), (b) σ (Gaussian resolution), (c) α (transition point) and (d) n (power law).
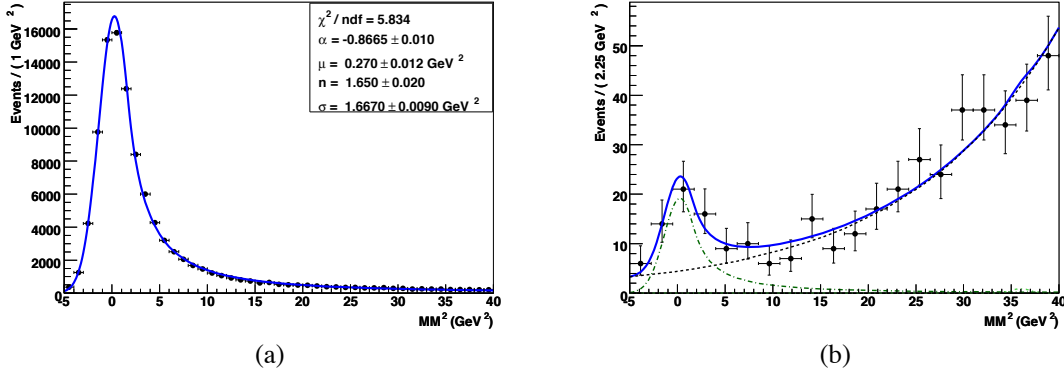
Figure 13: Fit to the distribution of the recoil mass squared m_X^2 for the events that pass (a) final selection w/o the IFR veto and (b) the complete final selection. All off-resonance events collected with the single photon trigger are shown. Solid blue line represents the total PDF, dot-dashed green line is for $e^+e^- \to \gamma\gamma$ contribution (fixed in the fit), and the dashed black like represents the continuum background.

This procedure assumes that for the photons that pass through the dead region in the EMC, the IFR detection efficiency is approximately constant in time.

To determine $N_{\gamma\gamma}^{\text{off}}$, we perform the yield fits to the off-resonance samples setting $N_{\text{signal}} = 0$. This is reasonable for the off-resonance data as well as the $\Upsilon(4S)$ and above-$\Upsilon(4S)$ data, since the branching fraction for the signal modes $\Upsilon(4S) \to \gamma A^0$ should be suppressed compared to $\Upsilon(3S)$ by over three orders of magnitude due to the large total width of $\Upsilon(4S)$. One example of such fit (for the merged off-resonance sample, taken with the single photon triggers and filters activated) is shown in Fig. 13. The value of ε_{IFR} for the various off-resonance samples is shown in Fig. 14. We compute the average IFR veto efficiency

$$\langle \varepsilon_{\text{IFR}} \rangle = (4.5 \pm 1.9) \times 10^{-4} \tag{4}$$

where the uncertainty covers the possible run-to-run spread in the values.

The advantage of this method of determining the peaking background from $e^+e^- \to \gamma\gamma$ events is that it is completely data-driven, and the samples satisfy all the standard selection criteria. Additionally, the yield $N_{\gamma\gamma}^{\text{off}}$ (and therefore, ε_{IFR}) includes any possible peaking background from ISR production of light resonances, such as ρ, ω, or ϕ. On the other hand, the uncertainty is limited by the off-resonance statistics. For the final publication of this analysis, we plan to measure the photon detection efficiency (including IFR veto) in alternative samples, such as radiative di-muons and bhabhas.

5.3 Fits on the "LowHigh" Datasets

We test the fitting procedure on the combination of `AllEvents-Run7-R24-Y3S-OnPeak-Low` and `AllEvents-Run7-R24-Y3S-OnPeak-High`, which together amount to $2482 \, \text{pb}^{-1}$, or 10.4 million $\Upsilon(3S)$. The sample in the m_X^2 interval $-5 < m_X^2 < 40 \, \text{GeV}^2$ (which corresponds to $E_\gamma^* > 3.25 \, \text{GeV}$ at $\Upsilon(3S)$) contains 82 events.

19
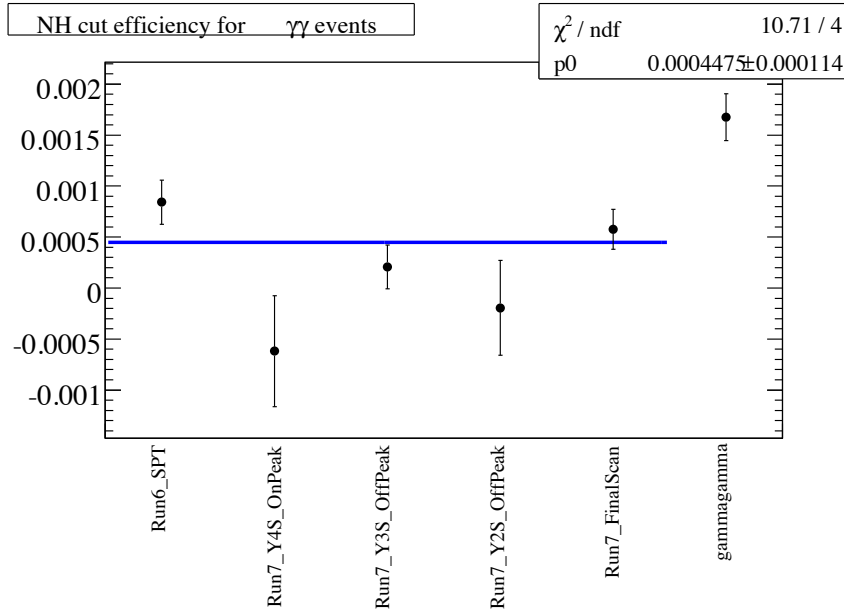
Figure 14: IFR veto efficiency $\varepsilon_{\mathrm{IFR}}$ for the various off-resonance samples and for the SP-10 $e^+e^- \to \gamma\gamma$ Monte Carlo (SP-1074). The Monte Carlo value is not used in the average.

Two examples of such fits are shown in Fig. 15. We fix the $e^+e^- \to \gamma\gamma$ contribution to

$$N_{\gamma\gamma} = 23998 \times (4.5 \pm 1.9) \times 10^{-4} = 10.4 \pm 4.4 \tag{5}$$

where $N^{\mathrm{on}}_{\gamma\gamma\,\mathrm{EMC}} = 23998$. The yields of the signal and continuum background events are floated, as well as the shape of the continuum background.

Fig. 16 shows the results of the fits at different values of m_{A^0}; the values of the fitted background parameters are plotted in Fig. 17. The signal yield is generally consistent with zero, with a few $2 - 3\sigma$ fluctuations. We show the statistical error and the total uncertainty in the yield, taking into account the dominant uncertainty due to the error on $\langle\varepsilon_{\mathrm{IFR}}\rangle$. For a 2.48 fb^{-1} sample the systematic error is not significant.

Taking into account signal efficiency, we can compute the expected upper limits on the branching ratio of $\Upsilon(3S) \to \gamma A^0$ for the full data sample of 120M $\Upsilon(3S)$ events. We assume that the statistical error at each point will increase as a square of the luminosity, and the systematic error due to uncertainty in $\langle\varepsilon_{\mathrm{IFR}}\rangle$ will scale linearly with the luminosity (since we have already used all available off-resonance statistics). The projected 90% C.L. upper limits for 120M $\Upsilon(3S)$ decays are shown in Fig. 18. We expect to reach the sensitivity of $(1 - 5) \times 10^{-6}$ for masses below $m_{A^0} < 6$ GeV, with the sensitivity at low values of mass (high photon energies) limited by the systematics on the $e^+e^- \to \gamma\gamma$ efficiency.
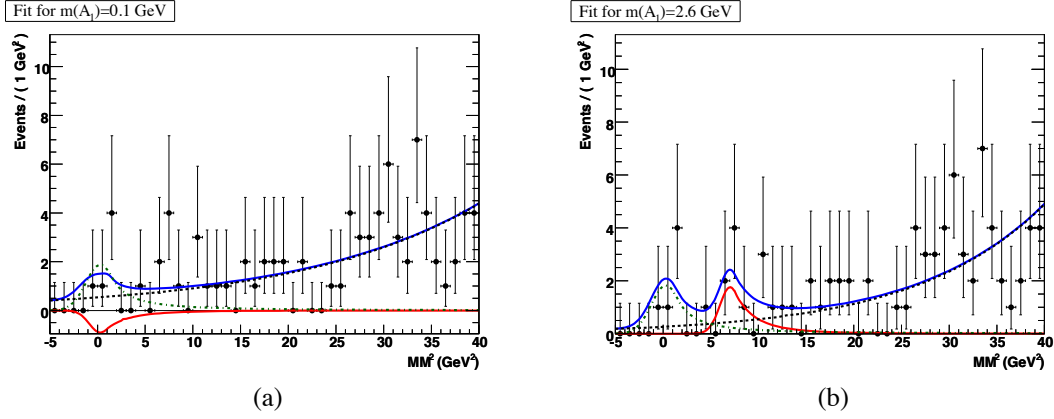
Figure 15: Two examples of the fits to the 2.48 fb^{-1}. (a): fit for $m_{A^0} = 0.1$ GeV, which returns $N_{\text{sig}} = -4.2^{+4.2}_{-3.4}$ events. (b): fit for fit for $m_{A^0} = 2.6$ GeV, which returns $N_{\text{sig}} = +7.7^{+4.5}_{-3.7}$ events. Solid blue line represents the total PDF, solid red line shows the signal PDF, dot-dashed green line is for $e^+e^- \to \gamma\gamma$ contribution (fixed in the fit), and the dashed black like represents the continuum background.
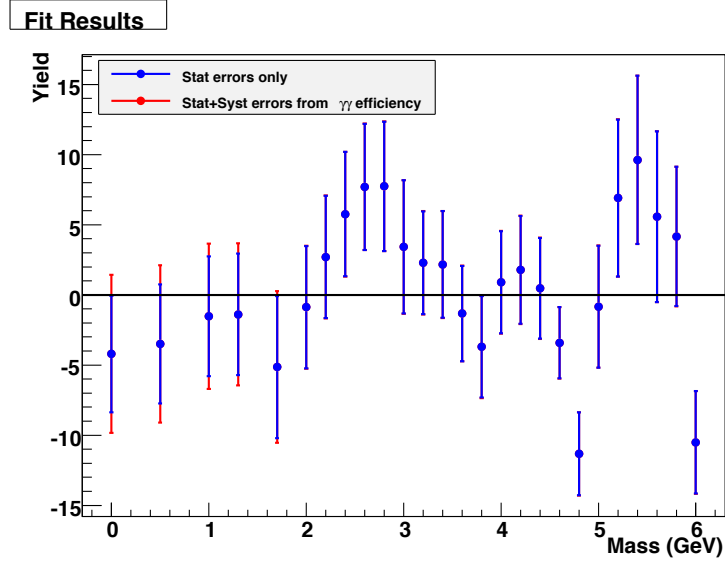
Figure 16: Signal yields N_{sig} as a function of assumed mass m_{A^0} for a sample of 2.48 fb^{-1} at $\Upsilon(3S)$. Blue error bars are statistical only, and the red error bars include the systematic contribution due to the uncertainty in $N^{on}_{\gamma\gamma}$.
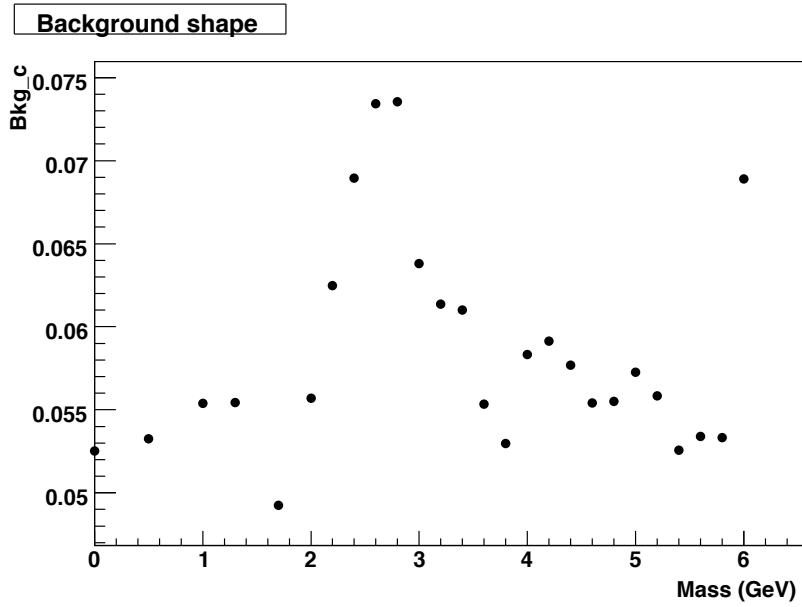
Figure 17: Exponential background parameter Bkg_c as a function of assumed mass m_{A^0} for a sample of $2.48\,\mathrm{fb}^{-1}$ at $\Upsilon(3S)$. Statistical uncertainties only.
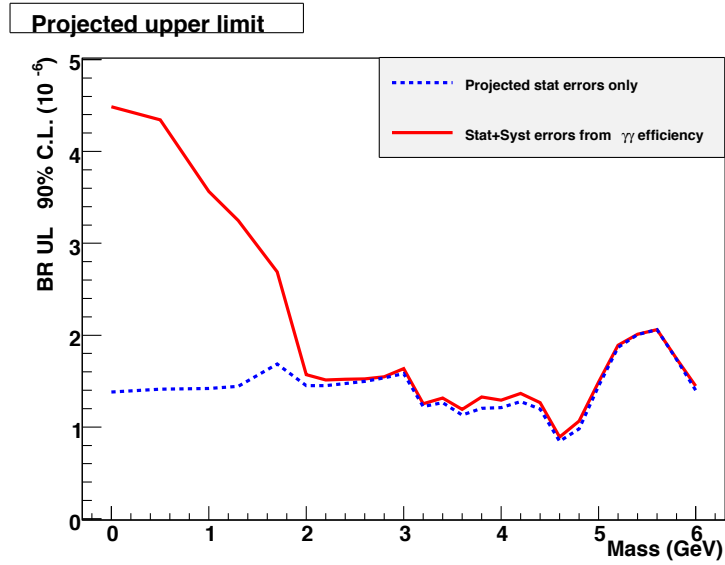
Figure 18: Expected 90% C.L. upper limits for the full Run7 $\Upsilon(3S)$ data sample. The dashed blue line shows the statistical uncertainties only, the solid red line includes the systematic uncertainties due to the error on $N_{\gamma\gamma}^{on}$.

6 Yield Fit in the Low Energy Region

The main backgrounds in the low energy region $1.5 < E_\gamma^* < 3.7$ GeV come from $e^+e^- \to \gamma\gamma$ process (above 3 GeV) and from the low-angle radiative bhabhas (which account for a very fast rise in the background below 3 GeV). In this region, the polar angle distribution of the background events peaks strongly in the forward and backward direction, which can provide additional separation between signal and background. Thus, we perform the fit in two dimensions $(m_X^2, \cos\theta_\gamma^*)$.

Our fitting strategy is as follows. We fit the distributions of $(m_X^2, \cos\theta_\gamma^*)$ in windows around the nominal assumed mass squared $m_{A^0}^2$. The windows are chosen to reduce the variation in the background levels at the ends of the internal, while keeping sufficient sidebands around the nominal mass. We typically take $5\,\mathrm{GeV}^2$ (5-10σ) below the nominal mass and $10\,\mathrm{GeV}^2$ (10-20σ) above the nominal mass. We fit the distributions to combination of the following PDFs:

(a) Signal PDF, determined from Monte Carlo, with floated signal yield. We use the same dependence on the assumed true mass of A^0 for the m_X^2 PDF as described in Section 5. We fit the distribution of $\cos\theta_\gamma^*$ to the second-order form $1 + p_2\cos^2\theta_\gamma^*$, with p_2 fixed to the value obtained in the Monte Carlo samples. The fitted values of p_2 are consistent with the expected value of 1, with a mild dependence on photon energy (see Fig. 19).

(b) "Continuum" background from radiative bhabhas and $e^+e^- \to \gamma\gamma$. This background is parameterized in m_X^2 by a smooth function (sum of a constant and exponential) with the shape parameters floated in the data fit. We use a 4-th order polynomial function with floated parameters for the $\cos^2\theta_\gamma^*$ PDF.

(c) We have investigated adding an explicit component for the $e^+e^- \to \gamma\gamma$ background, with the PDFs determined from the `SP-1074` Monte Carlo sample. This does not change the performance of the fitter significantly, so we reserve this option for systematic studies.

6.1 Fits on the "BGFilterSkim-High" Dataset

We test the fitting procedure on the `BGFilterSkim` dataset from the end of February 2008, which amounts to $2393\,\mathrm{pb}^{-1}$, or 10.0 million $\Upsilon(3S)$. The sample in the m_X^2 interval $30 < m_X^2 < 70\,\mathrm{GeV}^2$ (which corresponds to $1.8 < E_\gamma^* < 3.7$ GeV at $\Upsilon(3S)$) and $|\cos\theta_\gamma^*| < 0.6$ contains 28,412 events. The spectrum is shown in Fig. 20. For comparison, we also overlay the spectrum for the off-resonance data.

Examples of such fits at a few different values of m_{A^0} are shown in Fig. 21-25. Fig. 26 summarizes the results; the values of the fitted background parameters are plotted in Fig. 27. The signal yield is 1-2σ away from zero for most data points, except for the point at $m_{A^0} = 6.5$ GeV, which shows a larger fluctuation. This is addressed below.

Taking into account signal efficiency, we can compute the expected upper limits on the branching ratio of $\Upsilon(3S) \to \gamma A^0$ for the full `BGFSingleGammaInvisibleLowE` data sample of 82M $\Upsilon(3S)$ events. We assume that the statistical error at each point will increase as a square of the luminosity, and that the systematic errors due to PDF uncertainties are small. The projected 90% C.L. upper limits for 82M $\Upsilon(3S)$ decays are shown in Fig. 28. We expect to reach the sensitivity of $(1 - 12) \times 10^{-6}$ for masses below $m_{A^0} < 8$ GeV.
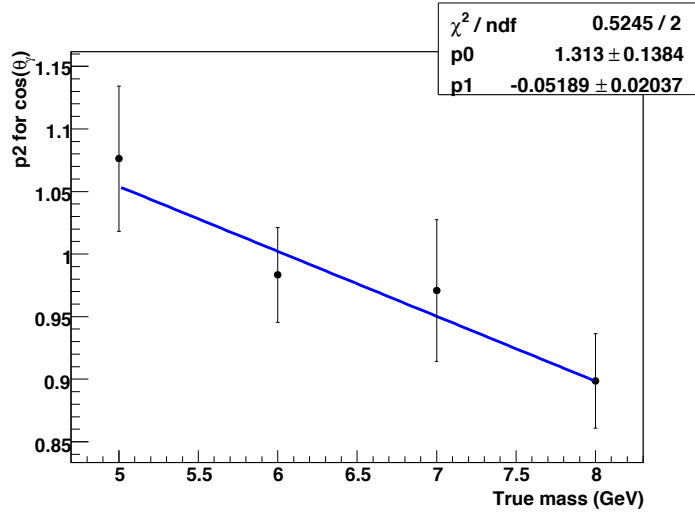
Figure 19: Dependence of the signal PDF parameter p_2 for the $\cos\theta_\gamma^*$ PDF on the true mass m_{A^0}.
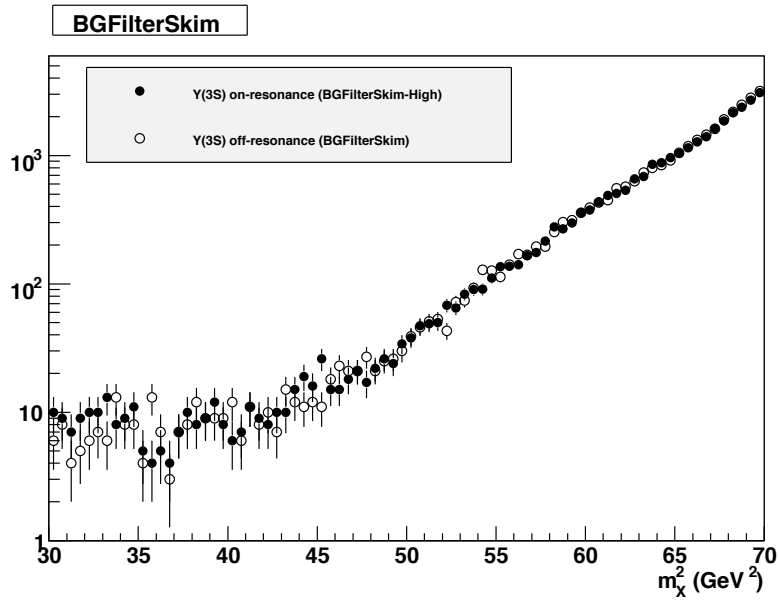

Figure 20: Spectrum of m_X^2 for events selected with `BGFSingleGammaInvisibleLowE` line. Solid points: $\Upsilon(3S)$ 2.4 fb^{-1} sample, open circles: off-$\Upsilon(3S)$ data (2.6 fb^{-1}).
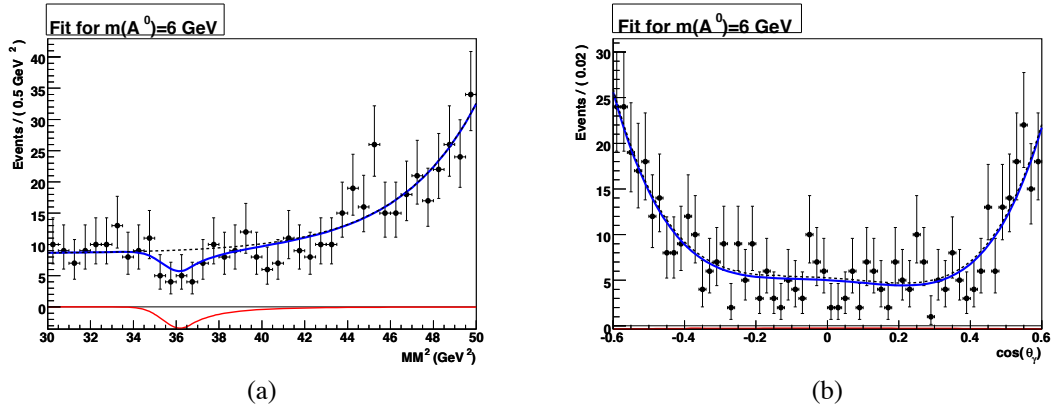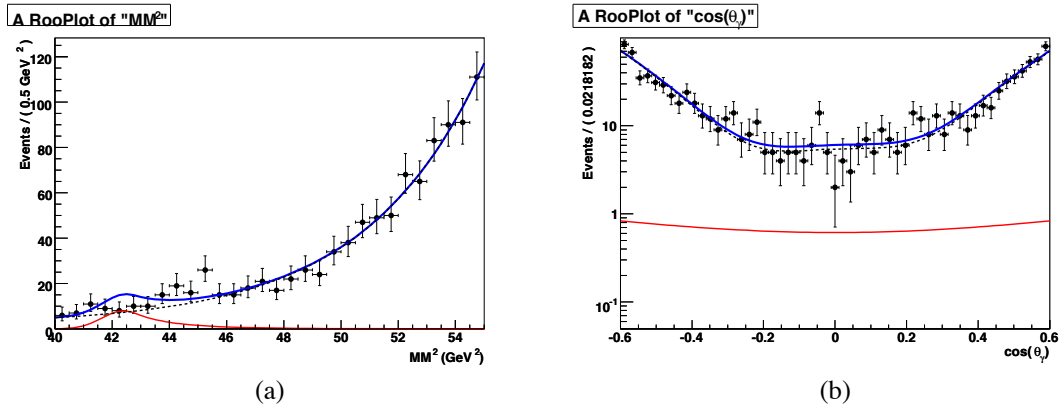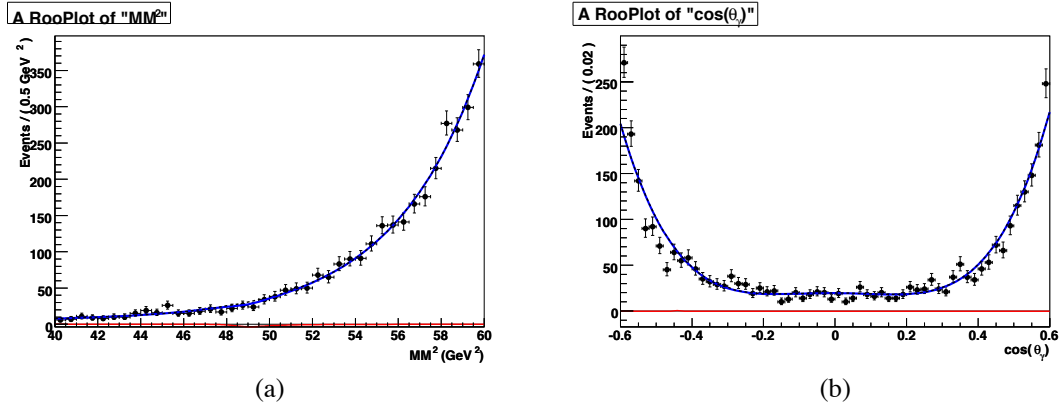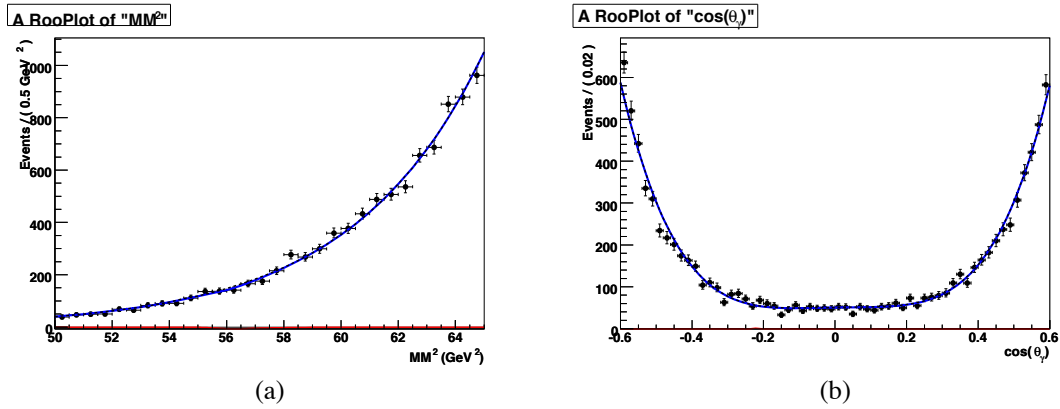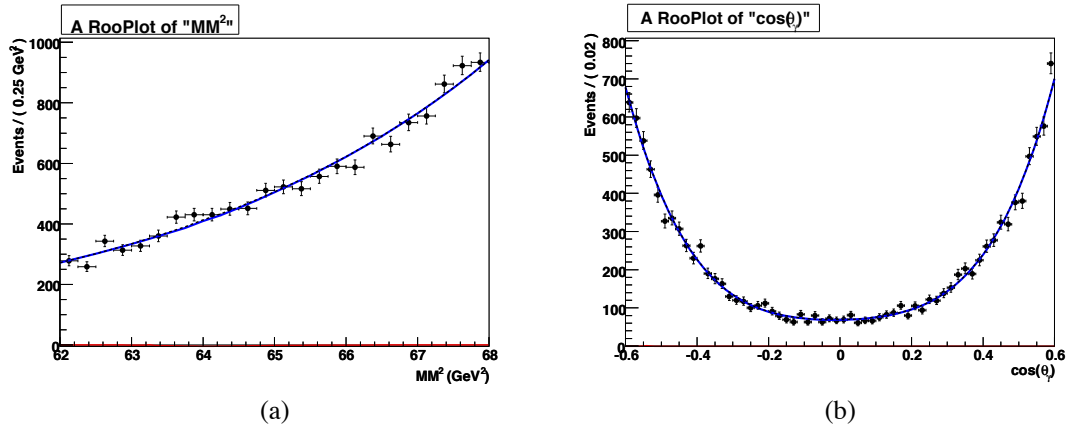
Figure 21: A sample fit to the 2.4 fb^{-1} sample for $m_{A^0} = 6.0$ GeV, which returns $N_{\text{sig}} = -20.0 \pm 8.2$ events. (a): projection onto m_X^2 and (b): projection onto $\cos\theta_\gamma^*$. The data points are in black (statistical errors only). Solid blue line represents the total PDF, solid red line shows the signal PDF, and the dashed black like represents the continuum background.

Figure 22: A sample fit to the 2.4 fb^{-1} sample for $m_{A^0} = 6.5$ GeV, which returns $N_{\text{sig}} = 42 \pm 13$ events. (a): projection onto m_X^2 and (b): projection onto $\cos\theta_\gamma^*$. The data points are in black (statistical errors only). Solid blue line represents the total PDF, solid red line shows the signal PDF, and the dashed black like represents the continuum background.

25

(a)

(b)

Figure 23: A sample fit to the 2.4 fb^{-1} sample for $m_{A^0} = 7.0$ GeV, which returns $N_{\text{sig}} = -17 \pm 12$ events. (a): projection onto m_X^2 and (b): projection onto $\cos\theta_\gamma^*$. The data points are in black (statistical errors only). Solid blue line represents the total PDF, solid red line shows the signal PDF, and the dashed black like represents the continuum background.

(a)

(b)

Figure 24: A sample fit to the 2.4 fb^{-1} sample for $m_{A^0} = 7.5$ GeV, which returns $N_{\text{sig}} = -29 \pm 29$ events. (a): projection onto m_X^2 and (b): projection onto $\cos\theta_\gamma^*$. The data points are in black (statistical errors only). Solid blue line represents the total PDF, solid red line shows the signal PDF, and the dashed black like represents the continuum background.

26

(a)   (b)

Figure 25: A sample fit to the 2.4 fb^{-1} sample for $m_{A^0} = 8.0$ GeV, which returns $N_{\text{sig}} = -172 \pm 72$ events. (a): projection onto m_X^2 and (b): projection onto $\cos\theta_\gamma^*$. The data points are in black (statistical errors only). Solid blue line represents the total PDF, solid red line shows the signal PDF, and the dashed black like represents the continuum background.
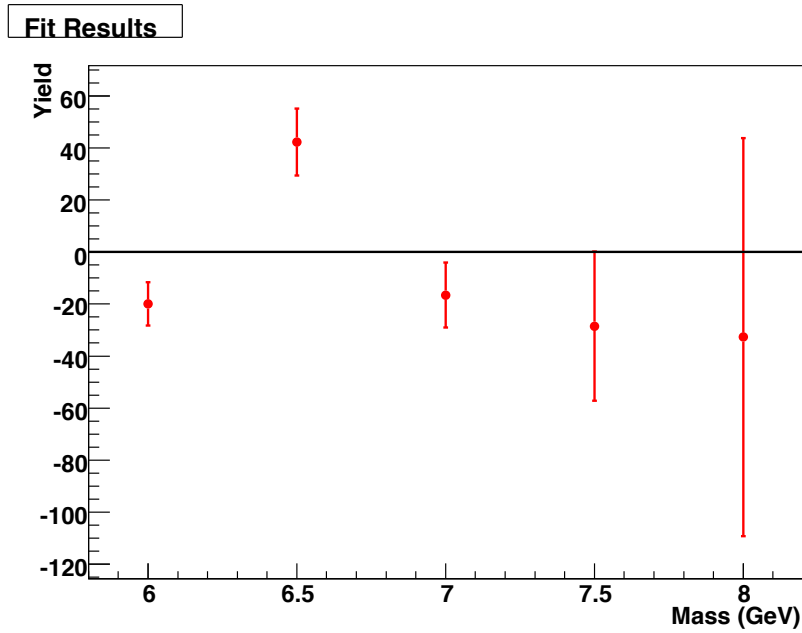
Figure 26: Signal yields N_{sig} as a function of assumed mass m_{A^0} for a sample of 2.4 fb^{-1} at $\Upsilon(3S)$. Statistical uncertainties only.

27
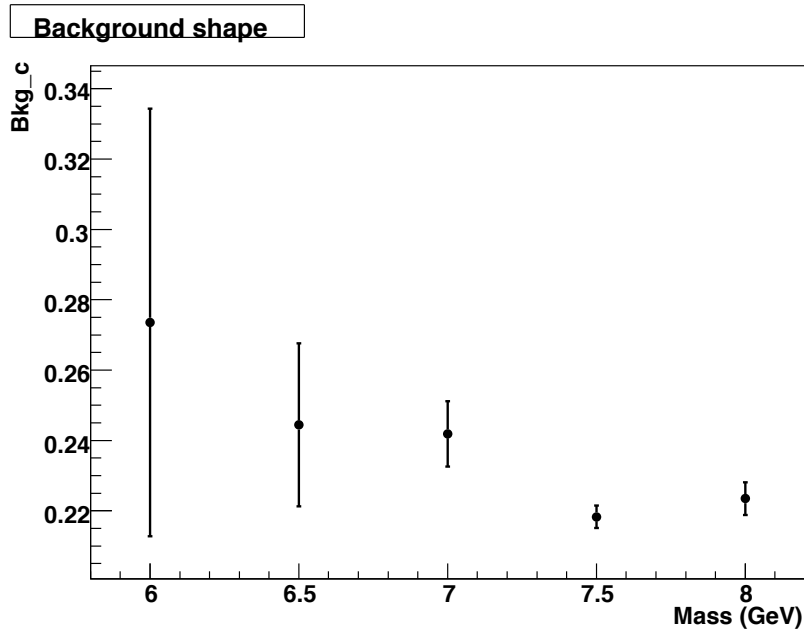
Figure 27: Exponential background parameter `Bkg_c` as a function of assumed mass m_{A^0} for a sample of 2.4 fb^{-1} at $\Upsilon(3S)$. Statistical uncertainties only.
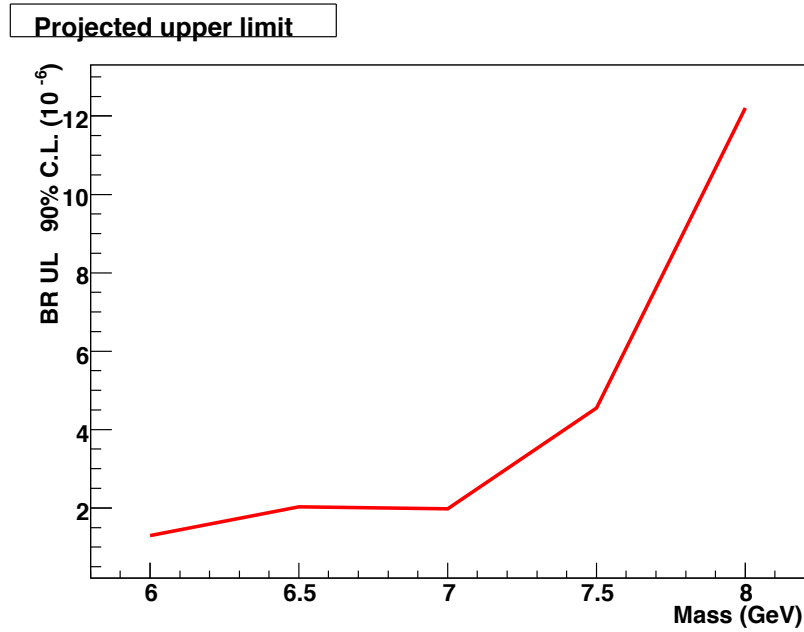
Figure 28: Expected 90% C.L. upper limits for the full Run7 $\Upsilon(3S)$ data sample with `BGFSingleGammaInvisibleLowE` line (statistical uncertainties only).
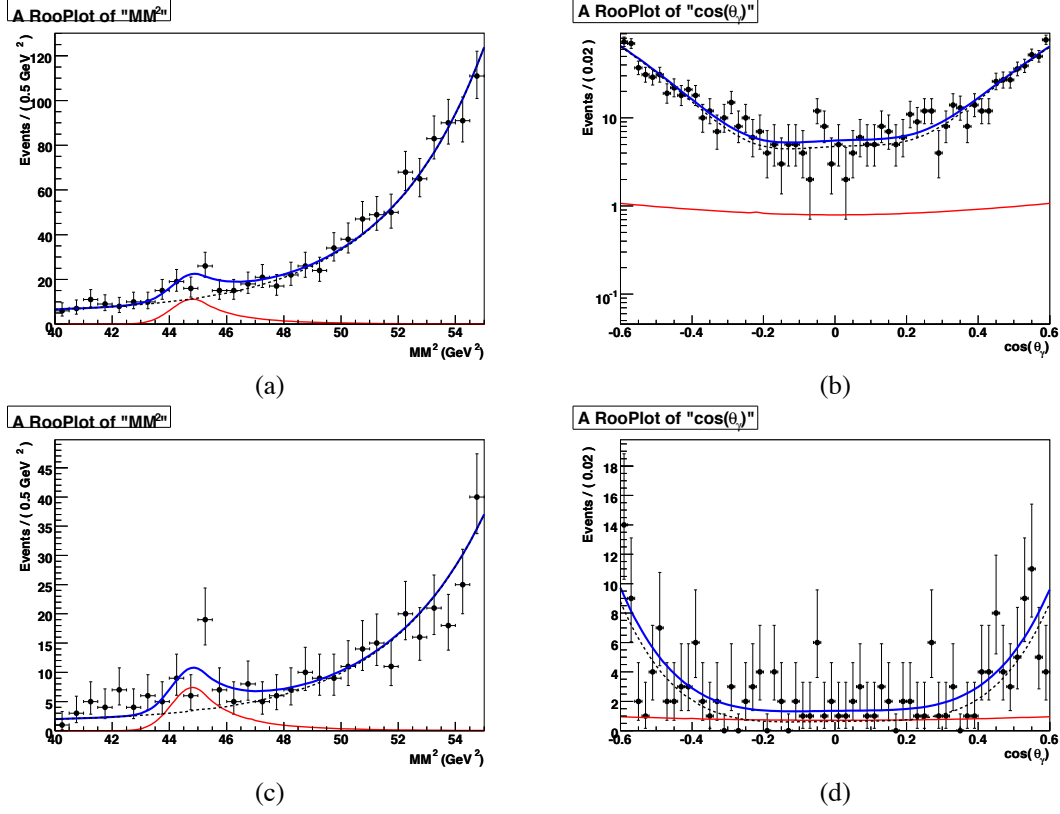
Figure 29: Fit to the 2.4 fb^{-1} sample for $m_{A^0} = 6.683$ GeV. (a,c): projections onto m_X^2 and (b,d): projection onto $\cos\theta_\gamma^*$. Plots (c)-(d) were made with a cut on the likelihood ratio $L_{\rm sig}/L_{\rm tot} > 0.5$, evaluated with the PDF in the orthogonal variable. The data points are in black (statistical errors only). Solid blue line represents the total PDF, solid red line shows the signal PDF, and the dashed black like represents the continuum background.

6.2 A Bump

Fig. 26 indicates some excess of events near $m_{A^0} = 6.5$ GeV. A closer look at that region finds a significant bump at the mass $m_{A^0} = 6.683 \pm 0.025$ GeV (obtained from a fit with floated m_{A^0}). The yield returned by the fit is $N_{\rm sig} = 57 \pm 13$, which corresponds to the branching ratio $\mathcal{B}(\Upsilon(3S) \to \gamma A^0) \times \mathcal{B}(A^0 \to \text{invisible}) = (20 \pm 5) \times 10^{-6}$ (statistical errors only). Taken at face value, *i.e.* by evaluating the log-likelihood ratio $\sqrt{2\ln(L_{\max}/L_0)}$, the significance of the excess is 5.1σ (here, naturally, $L_{\max}$ is the maximum likelihood with the floated signal, and L_0 is the likelihood for the fit with the signal yield fixed at zero). Fig. 29 shows the projections from the fit for the fixed $m_{A^0} = 6.683$ GeV.

To investigate this peak further, we have looked at the distribution of m_X^2 in the off-$\Upsilon(3S)$ and off-$\Upsilon(2S)$ data, as well as in a sample of 1.30 fb^{-1} of Run7 on-$\Upsilon(2S)$ data. The spectrum near $m_X^2 = 45$ GeV^2 is shown in Fig. 30. There does appear to be a visible excess of events in the on-$\Upsilon(3S)$ data compared to the off-resonance data, but this excess is not shared by the
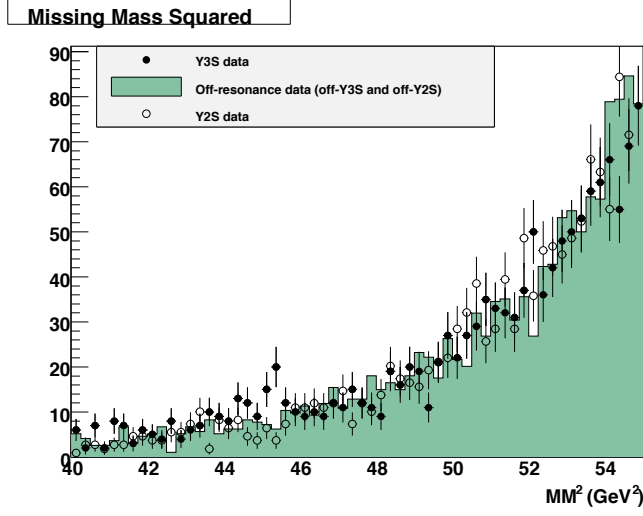
Figure 30: Spectrum of the recoil mass in the region $40 < m_X^2 < 55\,\mathrm{GeV}^2$. The black circles points show events selected in a 2.4 fb^{-1} sample on $\Upsilon(3S)$. The open circles show the 1.30 fb^{-1} sample collected on $\Upsilon(2S)$. The teal-shaded histogram shows the sum of the spectra collected off-$\Upsilon(3S)$ and off-$\Upsilon(2S)$ (scaled to the same number of events as the on-$\Upsilon(3S)$ histogram).
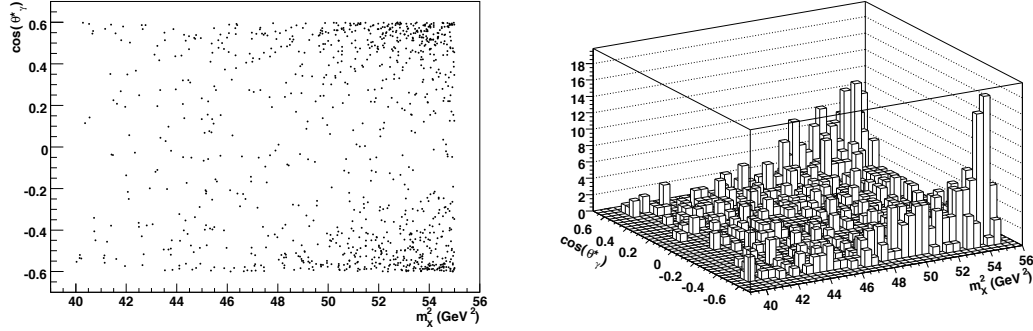
Figure 31: Distribution of $(m_X^2, \cos\theta_\gamma)$ for the selected events in the interval $40 < m_X^2 < 55\,\mathrm{GeV}^2$.

on-$\Upsilon(2S)$ data.[3]

We have also checked that the 2d distribution of $(m_X^2, \cos\theta_\gamma)$ does not show strong correlations between the two variables. There do appear to be two components in the data (besides the signal): $e^+e^- \to \gamma\gamma$ distribiton which is shallow in both variables, and radiative bhabhas, which turn on around $m_X^2 = 48\,\mathrm{GeV}^2$ and have a highly peaked distribution in $\cos\theta_\gamma$ (Fig. 31).

We have looked at potential pathologies in $\Upsilon(3S)$ data, but could not find any problems that

---

[3]If the peak were a real $\Upsilon \to \gamma A^0$ decay, we would have expected roughly the same branching ratio for $\Upsilon(2S)$ and $\Upsilon(3S)$, so the 1.3 fb^{-1} sample on $\Upsilon(2S)$ should have yielded roughly the same number of events as the 2.4 fb^{-1} sample at $\Upsilon(3S)$. Such estimates are probably good to within 50% or so, without any $\Upsilon(2S)$ Monte Carlo samples.

would not show up in the off-resonance or $\Upsilon(2S)$ data. For example, the angular distribution of the events (in both CMS and Lab) for the region $42 < m_X^2 < 48$ GeV^2 is shown in Fig. 32, and the distribution of run numbers is shown in Fig. 33. We also show the $N-1$ selection plots in Fig. 75-76 in Appendix A.
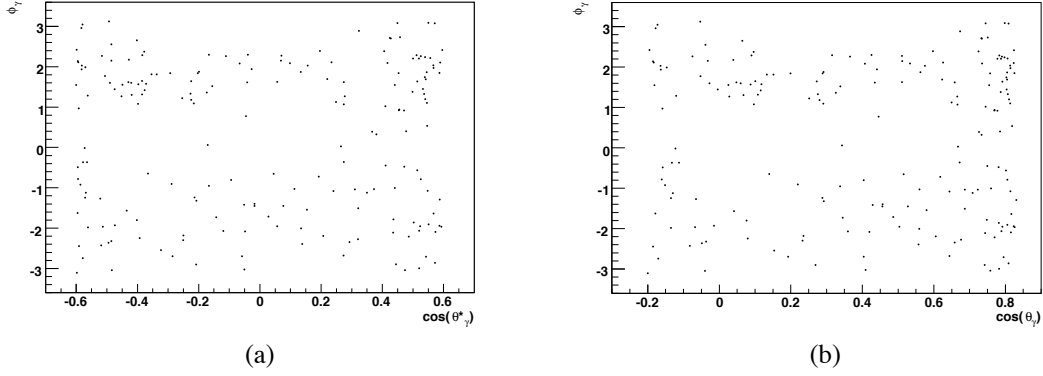
Figure 32: Distribution of $(\cos\theta_\gamma, \phi_\gamma)$ for the selected events in the interval $42 < m_X^2 < 48$ GeV^2. (a): CMS, (b): Lab.
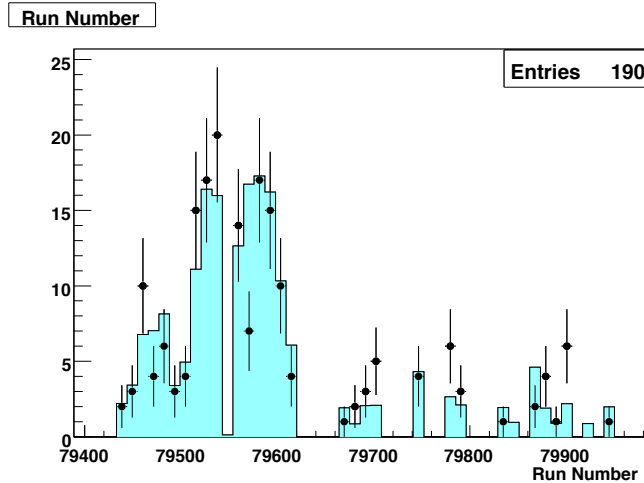
Figure 33: Run numbers for the selected events in the ''BGFilterSkim-High'' sample. Data points show the events in the interval $42 < m_X^2 < 48$ GeV^2, and the teal histogram shows the full data sample (scaled to the same integral).

The recoil mass $m_X = 6.68$ GeV corresponds to the photon energy of $E_\gamma^* = 3.02$ GeV. We have investigated if the peak could be a side-effect of some cut around $E_\gamma^* = 3$ GeV. The only cut at that value is the requirement $E_\gamma^* \geq 3$ GeV in the BGFSingleGammaInvisible line, though the BGFilter requirements are applied to EmcClusters with imprecise energy calibration, and the actual cut corresponds to 3.18 GeV (see Fig. 34). Nevertheless, we have investi-

31

e1Mag {BGFSingleGammaInvisible&&e1Mag<3.5}
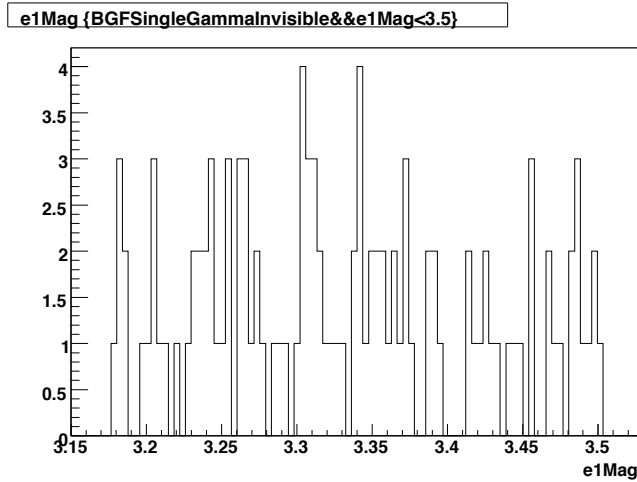
Figure 34: Energy distribution E_γ^* for events passing `BGFSingleGammaInvisible` selection near threshold. The threshold appears at $E_\gamma^* = 3.18$ GeV.

gated if the peak would disappear if `BGFSingleGammaInvisible` threshold was lowered to $E_\gamma^* \geq 2.3$ GeV. We have reprocessed 122 runs that correspond to the `BGFilterSkim-High` dataset[4] with the lowered threshold. Three runs, 79559, 79609, and 79789, have failed in the reprocessing, and were omitted from the dataset. Fig. 35 compares the spectra from the original `BGFilterSkim-High` dataset (with the 3 runs removed) and the new `BGFilterReskim` dataset. In the region of interest, the spectra are identical[5], so the peak is not an artifact of BGFilter selection.

We have also ran the fit on the `BGFilterSkim-OffPeak` dataset (2.6 fb^{-1}). The results are shown in Fig. 36 for the representative data points. The fit at $m_{A^0} = 6.68$ GeV yields $N_{\text{sig}} = 1\pm11$ events. A finer scan with floated m_{A^0} found a local maximum at $m_{A^0} = 7.365\pm0.013$ GeV, which corresponds to $N_{\text{sig}} = 107 \pm 30$ events and $\sqrt{2\ln(L_{\max}/L_0)} = 3.9$ (Fig. 37). According to Section 7.3, this corresponds to the background fluctuation probability of 1.2×10^{-3} (3σ significance).

Even though the naïve significance calculation claims a 5σ effect, it is still not inconceivable to see a fluctuation at that level when one looks at the *entire* $0 < m_{A^0} < 8$ GeV mass range without any *a-priori* knowledge of the Higgs mass. Evaluating how likely such a fluctuation is requires detailed toy Monte Carlo simulation (see Section 7.3). An alternative way of evaluating the real significance of this peak would be to look at several other data samples of similar statistics *fixing* the value of m_{A^0} to the value found in the 2.4 fb^{-1} sample (or, in other words, to unblind the entire $\Upsilon(3S)$ dataset) [12]. This will happen in due time, after the appropriate review of this document.

---

[4]Many thanks to the computing management for providing me access to the dedicated CPU farm and pre-staging the XTC files.

[5]They differ a little near the ends of the interval $40 < m_X^2 < 55$ GeV^2, presumably due to changes of the calibration constants.
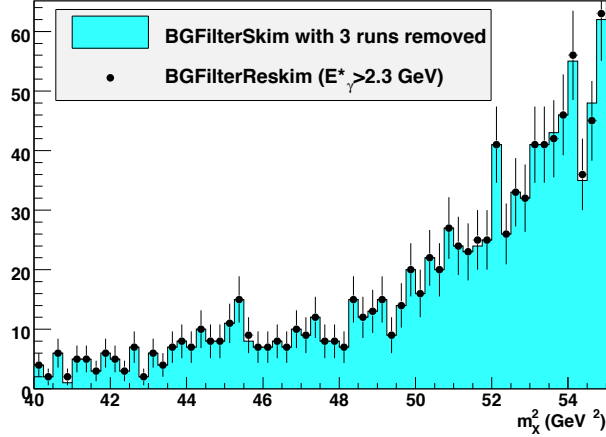
Figure 35: Distribution of the recoil masses m_X^2 near $m_X = 6.68$ GeV for the original `BGFilterSkim-High` dataset (filled cyan histogram) and the reprocessed `BGFilterReskim` dataset (solid points).

6.3 Options for Unblinding

Before unblinding the entire Run7 $\Upsilon(3S)$ dataset, we have looked at the performance of the fitter in various configurations, and the effect on the signal at $m_{A^0} = 6.68$ GeV. We summarize these studies below.

6.3.1 1d Fit

As an alternative, we checked the fit results for a 1d fit to the m_x^2 distribution only, in the different windows of $\cos\theta_\gamma^*$. The results are summarized in Table 3. The best signal significance is achieved for $|\cos\theta_\gamma^*| < 0.5$, which is also a result of blind optimization. The projection plot for the 1d fit with $|\cos\theta_\gamma^*| < 0.5$ is shown in Fig. 38.

Table 3: Selection efficiencies and results of the 1d fit for $m_{A^0} = 6.68$ GeV in the region $40 < m_X^2 < 55$ GeV^2 with different cuts on $\cos\theta_\gamma^*$.

Cut	Signal efficiency	Fitted Yield	Background	BR ($\times 10^{-6}$)	Significance (σ)
$\|\cos\theta_\gamma^*\| < 0.4$	0.166	29.9 ± 10.4	282 ± 19	18.0 ± 6.3	3.1
$\|\cos\theta_\gamma^*\| < 0.45$	0.188	32.6 ± 11.5	366 ± 22	17.3 ± 6.1	3.1
$\|\cos\theta_\gamma^*\| < 0.5$	0.212	39.1 ± 12.6	488 ± 25	18.4 ± 5.9	3.6
$\|\cos\theta_\gamma^*\| < 0.55$	0.238	38.7 ± 13.9	671 ± 29	16.2 ± 5.8	3.0
$\|\cos\theta_\gamma^*\| < 0.6$	0.264	36.1 ± 15.4	985 ± 34	13.7 ± 5.8	2.5

33
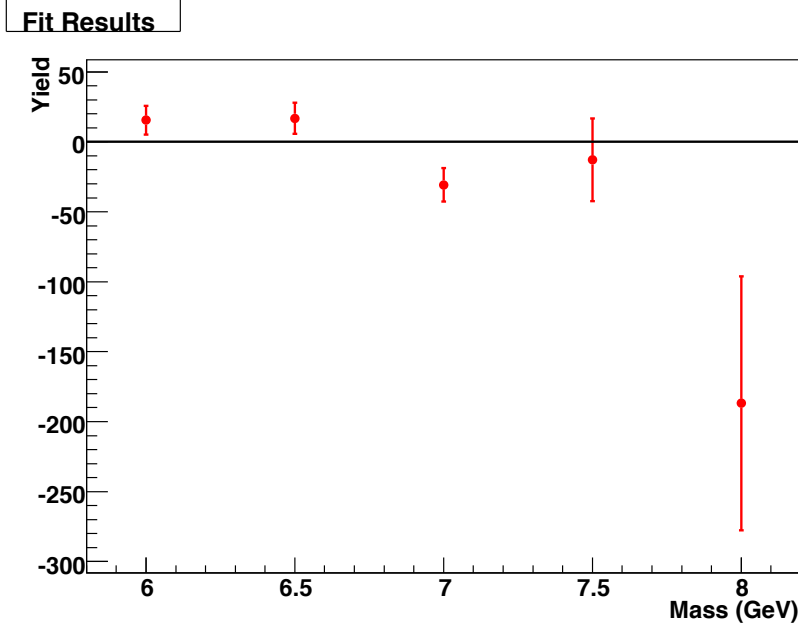
Figure 36: Signal yields N_{sig} as a function of assumed mass m_{A^0} for a sample of 2.6 fb^{-1} below $\Upsilon(3S)$. Statistical uncertainties only.

6.4 2d Fit with Additional Degrees of Freedom

We have investigating adding more degrees of freedom to the fit to improve the agreement between data and the PDF, especially below the $m_{A^0} = 6.68$ GeV peak. We have tried the following:

- Adding additional polynomial parameters to the background model

- Parameterizing the background PDF in m_X^2 as a function of $\cos\theta * *_\theta$

- Adding explicit PDFs for the main physics modes, $e^+e^- \rightarrow \gamma\gamma$ and radiative bhabhas

The best option, in terms of convergence and the overall χ^2, seems to be the last option. We restrict the range of $\cos\theta^*_\gamma$ to $|\cos\theta^*_\gamma| < 0.55$ (this is done to reduce potential peaking backgrounds, and improve the quality and convergence of the fit). We parameterize the background events as two explicit components, each with its own independent yield:

(a) $e^+e^- \rightarrow \gamma\gamma$: a sum of two PDFs: continuum and "peaking".

- The *peaking* component appears in the forward and backward region, presumably due to conversions in the forward support tube which miss the tracking volumes entirely but deposit a cluster in the calorimeter. We characterize it as a Gaussian in m_X^2 with $\mu = (47.4 \pm 0.3)$ GeV^2 and $\sigma = (3.3 \pm 0.2)$ GeV^2. These parameters are determined from the $e^+e^- \rightarrow \gamma\gamma$ Monte Carlo. The PDF in $\cos\theta^*_\gamma$ is a 4th order polynomial.

34

(a)　　　　　　　　　　　　　　(b)
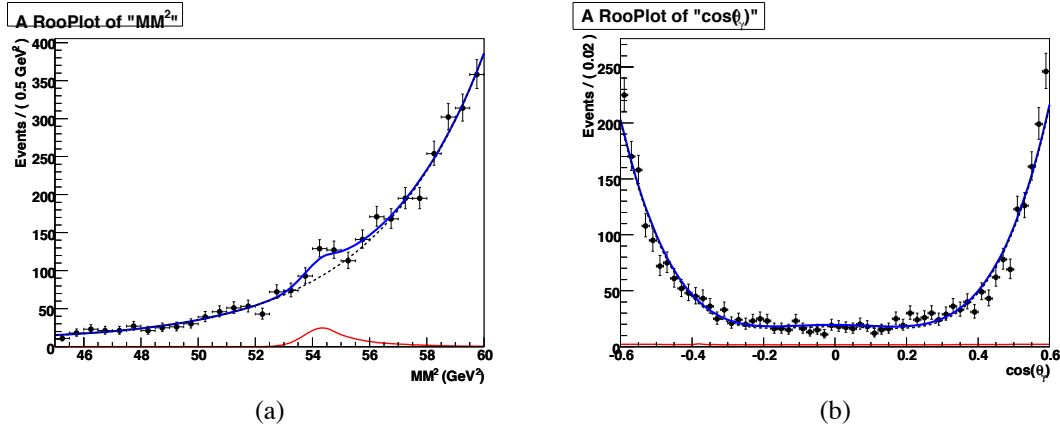
Figure 37: A sample fit to the 2.6 fb^{-1} off-resonance (below $\Upsilon(3S)$) sample for $m_{A^0} = 7.365 \pm 0.013$ GeV, which returns $N_{\text{sig}} = 107 \pm 30$ events with 3σ true statistical significance. (a): projection onto m_X^2 and (b): projection onto $\cos\theta_\gamma^*$. The data points are in black (statistical errors only). Solid blue line represents the total PDF, solid red line shows the signal PDF, and the dashed black like represents the continuum background.

- *Continuum* component is a 1st order polynomial in m_X^2 and a 2nd order polynomial in $\cos\theta_\gamma^*$.

The $e^+e^- \to \gamma\gamma$ PDF (from the Monte Carlo fit) is shown in Fig. 39.

(a) radiative bhabha component: a sum of two PDFs: continuum and "peaking".

- There is a possible *peaking* component in the radiative bhabha Monte Carlo near $m_X^2 = 50$ GeV^2. It is about 3σ significant, and it is too narrow compared to the signal PDF (Fig. 40) The off-resonance and on-resonance data do not support this component (the fraction is consistent with zero). It is most likely a fluctuation, but to be thorough, we include a Gaussian component in the radiative bhabha PDF (we have also requested 10x more Monte Carlo events). We fix the PDF parameters to the Monte Carlo values, and float the fraction in the data fit. The PDF in $\cos\theta_\gamma^*$ is a 4th order polynomial with all but the 4th coefficient zero.

- *Continuum* component is an exponential in m_X^2 and a 4th order polynomial in $\cos\theta_\gamma^*$.

The signal is parameterized as described in Section 6. We have also tried adding a 4th component to the fit (for the generic two-photon background), but since the fit had trouble converging (the distributions are presumably very close to either $e^+e^- \to \gamma\gamma$ or radiative bhabhas, so floated parameters very highly correlated).

In the data fit, we float 11 parameters:

- 4 polynomial coefficients for the bhabha distribution in $\cos\theta_\gamma^*$. A few of them are consistent with zero, so we could reduce the number of degrees of freedom if the fit to the full data set supports it.

- Exponential slope for the bhabha m_X^2 PDF

35
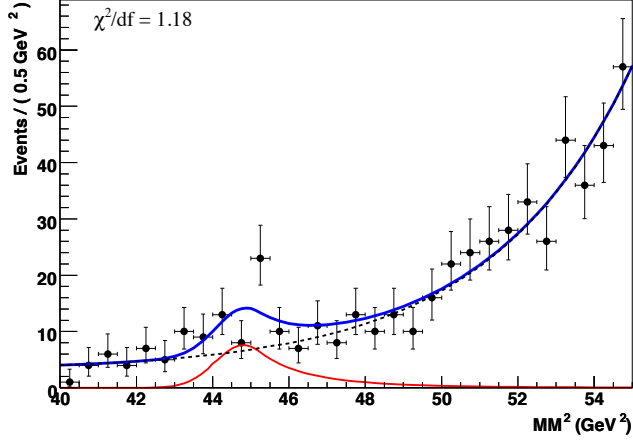
Figure 38: A 1d fit to the 2.4 fb^{-1} $\Upsilon(3S)$ sample for $m_{A^0} = 6.68$ GeV, with a cut $|\cos\theta^*_\gamma| < 0.5$. The data points are in black (statistical errors only). Solid blue line represents the total PDF, solid red line shows the signal PDF, and the dashed black like represents the continuum background.
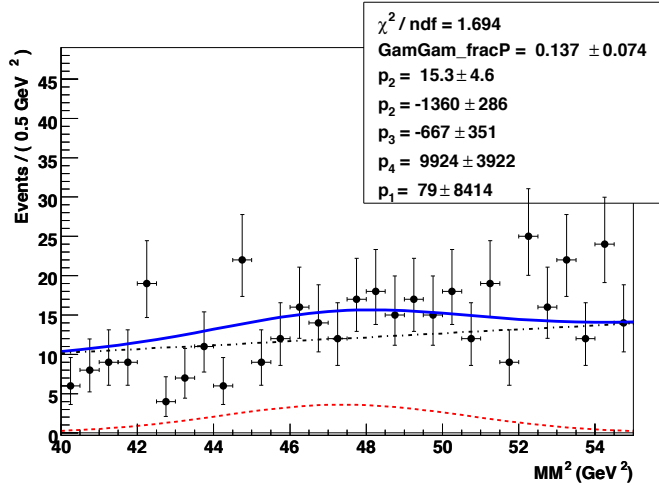
Figure 39: $e^+e^- \to \gamma\gamma$ PDF fit to the Monte Carlo sample

- Fraction of the peaking component in the bhabha PDF
- Fraction of the peaking component in the $e^+e^- \to \gamma\gamma$ PDF
- Slope parameter for the continuum component of $e^+e^- \to \gamma\gamma$ m^2_X PDF
- Yields for the signal, $e^+e^- \to \gamma\gamma$, and bhabha components.

First, we test this fit on the 2.6 fb^{-1} off-$\Upsilon(3S)$ sample. The results are summarized in Table 4 and Fig. 41. The signal yield is consistent with zero, and so is the peaking component of the
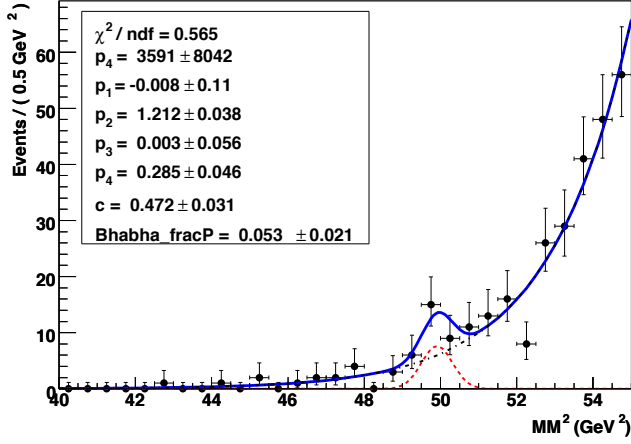
Figure 40: Bhabha PDF fit to the Monte Carlo sample

radiative bhabha background. The total PDF adequately describes the data.

Table 4: Results of the modified fit to the 2.6 fb^{-1} off-$\Upsilon(3S)$ sample. All floated parameters are shown. Statistical errors only.

PDF	Parameter	Value	Global corr
Bhabha $\cos\theta^*_\gamma$	p_1	0.12 ± 0.10	0.77
	p_2	0.95 ± 0.06	0.69
	p_3	-0.02 ± 0.06	0.77
	p_4	0.25 ± 0.06	0.67
Bhabha m_X^2	exp slope	0.40 ± 0.05	0.80
	peak frac	0.029 ± 0.018	0.46
$e^+e^- \to \gamma\gamma\, m_X^2$	peak frac	0.073 ± 0.037	0.48
	slope	$1.5 \times 10^2 \pm 1.2 \times 10^4$	0.15
	N_{bhabha}	490 ± 40	0.77
	$N_{\gamma\gamma}$	252 ± 40	0.84
	N_{sig}	0.6 ± 10.5	0.62
$-\ln(L)$		-2514.93	

The results of the modified fit to the 2.4 fb^{-1} on-resonance $\Upsilon(3S)$ sampke are summarized in Table 5 and Fig. 42. The signal yield significance is $\sqrt{2\ln(L_{\max}/L_0)} = 4.3\sigma$. The minimum value of the log-likelihood ($-\ln(L) = -2317.47$) agrees very well with the expectations from the toy Monte Carlo (Fig. 47), which indicates no major problems with the fit.
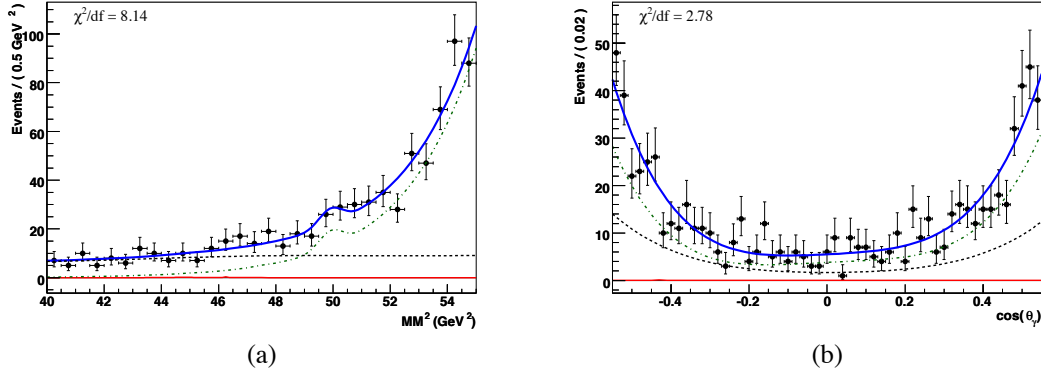
Figure 41: Results of the modified fit to the 2.6 fb^{-1} off-resonance (below $\Upsilon(3S)$) sample for $m_{A^0} = 6.68$ GeV. (a): projection onto m_X^2 and (b): projection onto $\cos\theta_\gamma^*$. The data points are in black (statistical errors only). Solid blue line represents the total PDF, solid red line shows the signal PDF, the dot-dashed green like represents the radiative bhabha background, and the black dashed line shows the $e^+e^- \to \gamma\gamma$ background.

Table 5: Results of the modified fit to the 2.4 fb^{-1} on-$\Upsilon(3S)$ sample. All floated parameters are shown. Statistical errors only.

PDF	Parameter	Value	Global corr
Bhabha $\cos\theta_\gamma^*$	p_1	0.07 ± 0.09	0.78
	p_2	1.03 ± 0.05	0.72
	p_3	0.02 ± 0.06	0.78
	p_4	0.23 ± 0.05	0.67
Bhabha m_X^2	exp slope	0.28 ± 0.03	0.80
	peak frac	0.00 ± 0.11	0.03
$e^+e^- \to \gamma\gamma\, m_X^2$	peak frac	0.00 ± 0.05	0.01
	slope	$3.5 \times 10^3 \pm 1.3 \times 10^4$	0.0
	N_{bhabha}	518 ± 47	0.84
	$N_{\gamma\gamma}$	145 ± 44	0.88
	N_{sig}	47 ± 13	0.50
$-\ln(L)$		-2317.47	

7 Fit Validation

We validate the fit using a large-statistics toy Monte Carlo samples. We study two issues: the bias of the fitter for various values of the true m_{A^0}, and the probability for a pure background sample to produce a given value of the computed significance, $\sqrt{2\ln(L_{\max}/L_0)}$.

For each toy experiment, we generate the background from the PDF obtained in the data fit, and embed a known amount of signal (0 or 10) events from the Monte Carlo simulations. For the high-energy region, we also embed 10 $e^+e^- \to \gamma\gamma$ events from the data, selected without the
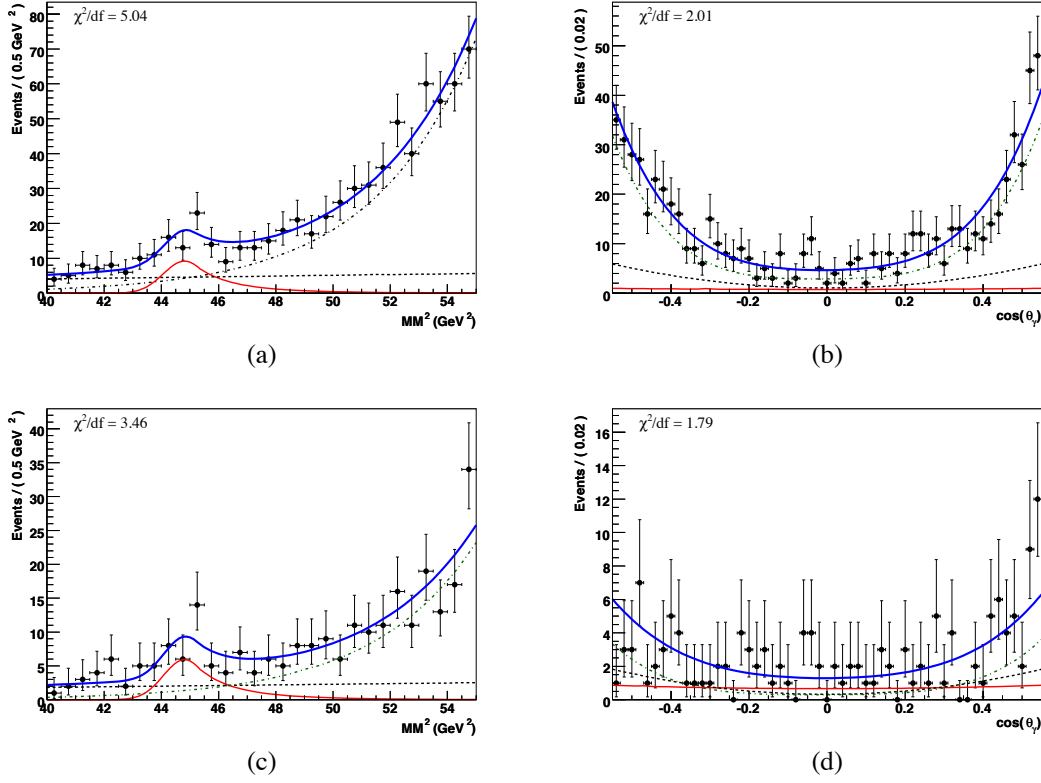
Figure 42: Results of the modified fit to the 2.4 fb^{-1} on-resonance sample for $m_{A^0} = 6.68$ GeV. (a,c): projections onto m_X^2 and (b,d): projection onto $\cos\theta_\gamma^*$. Plots (c)-(d) were made with a cut on the likelihood ratio $L_{\text{sig}}/L_{\text{tot}} > 0.5$, evaluated with the PDF in the orthogonal variable. The data points are in black (statistical errors only). Solid blue line represents the total PDF, solid red line shows the signal PDF, the dot-dashed green like represents the radiative bhabha background, and the black dashed line shows the $e^+e^- \rightarrow \gamma\gamma$ background.

IFR veto. The number of events in each fit is Poisson-distributed around the value observed in the data.

7.1 High energy region

The fit residual, pull mean, and pull RMS for each value of the true $m_{A^0} \le 6$ GeV are shown in Fig. 43-44 for $N_{\text{sig}}^{\text{true}} = 0$, and Fig. 45-46 for $N_{\text{sig}}^{\text{true}} = 10$. The bias is small, and the pull RMS is consistent with unity.

More detailed plots are shown in Appendix C.1.

7.2 Low Energy Region

The validation of the modified fit for $m_{A^0} = 6.68$ GeV is shown in Fig. 47. We draw 50 signal events from the PDF (Monte Carlo request for this mass has been submitted), and we also gener-
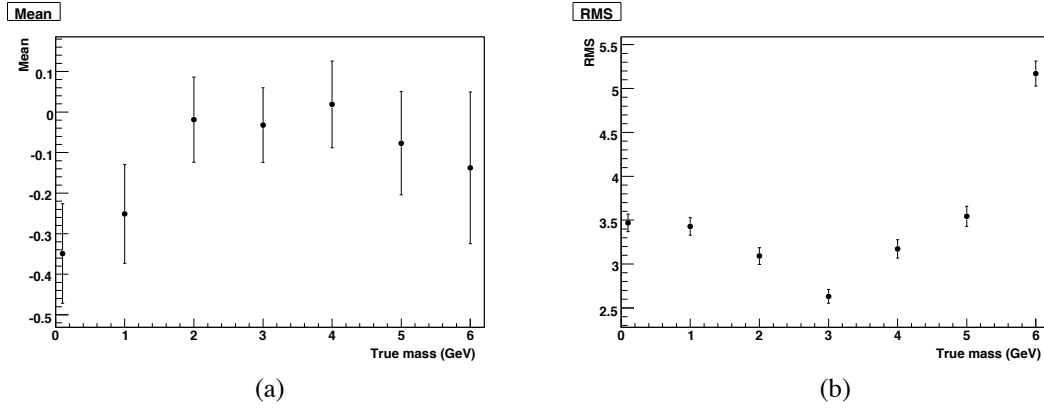
Figure 43: The average bias (a) and RMS (b) for the signal yield in the toy Monte Carlo results for $N_{\mathrm{sig}} = 0$ and $m_{A^0} \leq 6\,\mathrm{GeV}$.
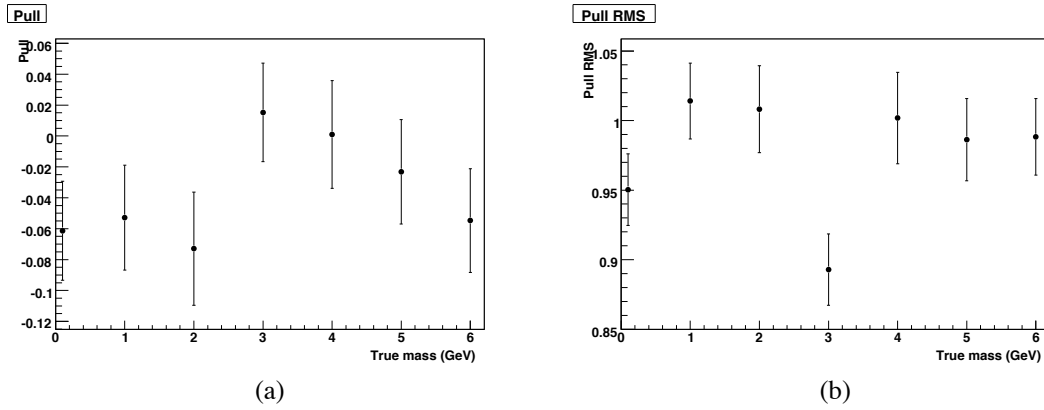
Figure 44: The average pull (a) and its RMS (b) for the signal events in the toy Monte Carlo results for $N_{\mathrm{sig}} = 0$ and $m_{A^0} \leq 6\,\mathrm{GeV}$.

ate the distributions of the $e^+e^- \to \gamma\gamma$ and bhabha backgrounds from the PDFs. No significant bias is observed, and the value of log-likelihood found in the fit to the $2.4\,\mathrm{fb}^{-1}$ on-$\Upsilon(3S)$ sample agrees well with the expectations from the toys.

7.3 Signal Significance

We also use the toy Monte Carlo to study the true signal significance, *i.e.* the probability for the pure background sample to fluctuate up to a given the value of the signal yield. To simplify normalization, we actually measure the fraction of toy experiments with a given value of the statistical significance measure $\mathcal{S} \equiv \Delta\chi^2 = 2\ln(L_{\mathrm{max}}/L_0)$ and $N_{\mathrm{sig}} > 0$. We perform the toy study in 3 mass ranges: $-5 < m_X^2 < 40\,\mathrm{GeV}^2$ ($m_{A^0} \leq 6\,\mathrm{GeV}$), $30 < m_X^2 < 60\,\mathrm{GeV}^2$ ($6 < m_{A^0} < 7.5\,\mathrm{GeV}$), and $60 < m_X^2 < 70\,\mathrm{GeV}^2$ ($7.5 < m_{A^0} < 8\,\mathrm{GeV}$). For each range, we generate the background events from the PDFs (1-d distributions for $m_{A^0} \leq 6\,\mathrm{GeV}$ and
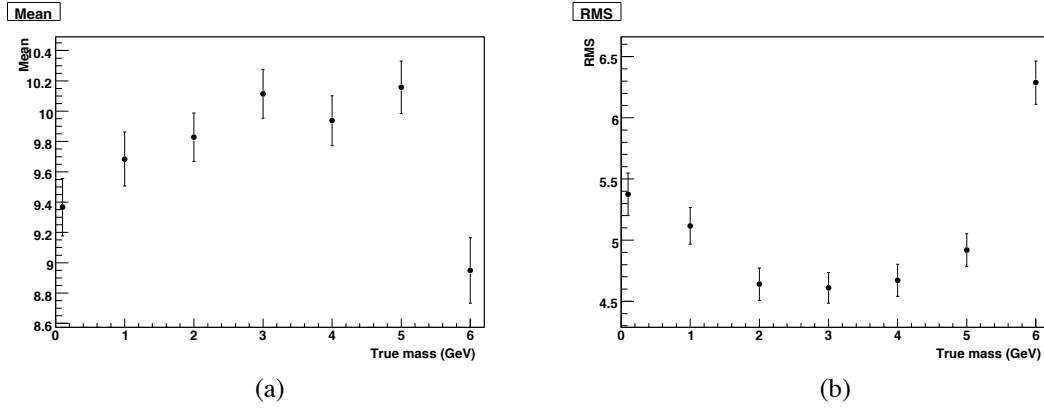
Figure 45: The average bias (a) and RMS (b) for the signal yield in the toy Monte Carlo results for $N_{\text{sig}} = 10$ and $m_{A^0} \leq 6$ GeV.
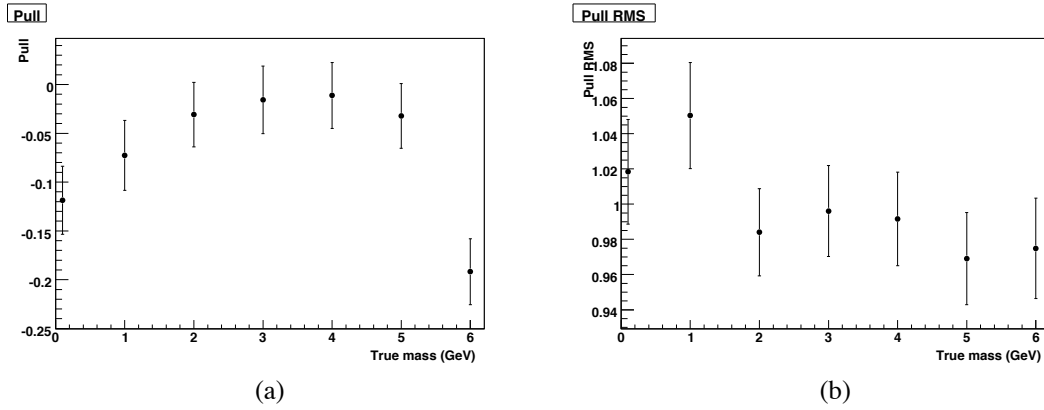
Figure 46: The average pull (a) and its RMS (b) for the signal events in the toy Monte Carlo results for $N_{\text{sig}} = 10$ and $m_{A^0} \leq 6$ GeV.

2-d distributions for $m_{A^0} > 6$ GeV), smearing the number of events according to the Poisson distributions around the value found in data. We run 5×10^6 toy experiments for $m_{A^0} \leq 6$ GeV range, and 10^6 toy experiments for $m_{A^0} > 6$ GeV. Unlike the data fits, we *float* the value of the assumed m_{A^0}. Thus, we measure the probability that a given value of $\mathcal{S}$ is observed *anywhere* in a given mass interval.

The plot of the fluctuation probability as a function of observed $\mathcal{S}$ for the high energy region ($m_{A^0} \leq 6$ GeV) is shown in Fig. 48. The plot for $6 < m_{A^0} < 7.5$ GeV is shown for Fig. 49. The probability is defined as

$$\mathcal{P}(\mathcal{S}_0) = \frac{N(\mathcal{S} \geq \mathcal{S}_0; N_{\text{sig}} > 0)}{N_{\text{gen}}}$$

where $N(\mathcal{S} \geq \mathcal{S}_0; N_{\text{sig}} > 0)$ is the number of events with the positive signal yield and the value of computed significance $\mathcal{S}$ above $\mathcal{S}_0$. We see that the probability for the background fluctuation behaves like a χ^2 distribution with 2-3 degrees of freedom (but closer to 2 d.f.). For instance, the

41
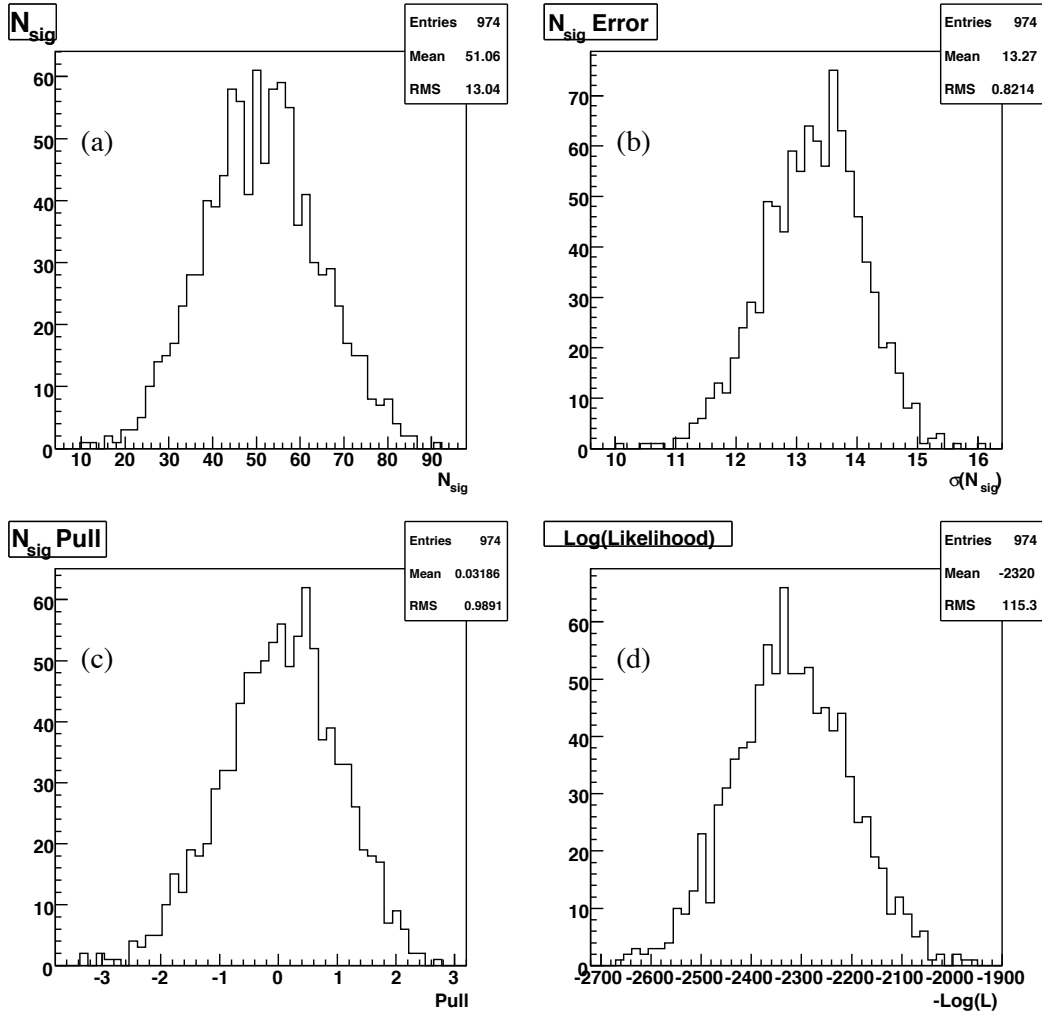
Figure 47: Toy Monte Carlo results for $N_{\text{sig}} = 50$ and $m_{A^0} = 6.7$ GeV. (a): signal yield, (b): error, (c): pull, (d): likelihood value $-\ln(L)$.

probability to observe $\mathcal{S} \geq 5.14^2$ in the interval $m_{A^0} \leq 6$ GeV is $(1.8 \pm 0.6) \times 10^{-6}$, and the probability to observe similar $\mathcal{S}$ in the range $6 < m_{A^0} < 7.5$ GeV is $(1 \pm 1) \times 10^{-6}$. The fits for the high-statistics range $7.5 < m_{A^0} < 8$ GeV are still running, but the probability could be estimated by scaling the result from the range $6 < m_{A^0} < 7.5$ GeV. Overall, we estimate that the probability to observe the effect of the same magnitude as seen in the $\Upsilon(3S)$ data ($\mathcal{S} \geq 5.14^2$) *anywhere* between $0 < m_{A^0} < 8$ GeV is $(3.1 \pm 1.4) \times 10^{-6}$, which corresponds to a 4.4-4.6σ effect.
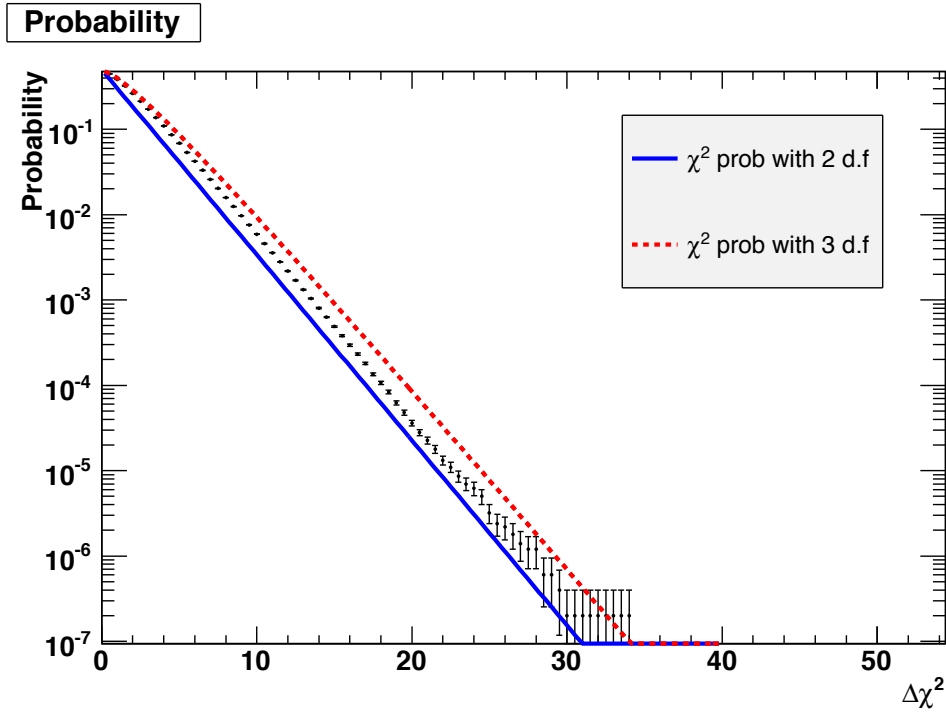
Figure 48: Probability to observe a background fluctuation with a given signal significance $\mathcal{S} \equiv \Delta\chi^2$ (plotted on the horizonal axis) anywhere in the range $m_{A^0} \leq 6$ GeV.

8 Systematics

The following systematic uncertainties need to be accounted for:

- *IFR veto efficiency* for the fixed $e^+e^- \rightarrow \gamma\gamma$ component of the fit in the High-Energy region is the dominant systematic uncertainty. We use the relative uncertainty of 42%, as described in Section 5.2.

- *PDF systematics*. In addition to PDF uncertainties determined by Monte Carlo statistics, the differences between the data and Monte Carlo are evaluated using the control $e^+e^- \rightarrow \gamma\gamma$ sample (see below).

- *Fit Bias* is determined from the Toy Monte Carlo studies, and is negligible (Section 7).

- *Selection efficiency*, determined from the Monte Carlo samples (see Fig. 9 and Fig. 11). To account for the spread in the values of the Monte Carlo efficiency (likely due to the oversampling of the background frames), we assign an efficiency uncertainty of 2%.

- *Gamma reconstruction efficiency* is provided by the neutrals group. At this point, no update is planned for the Summer 2008. We use a conservative 2% estimate for systematics.

- *Trigger systematics* is evaluated using a sample of events selected by `DigiFL3Prescale` (prescaled L3 pass-thru events)(see below).
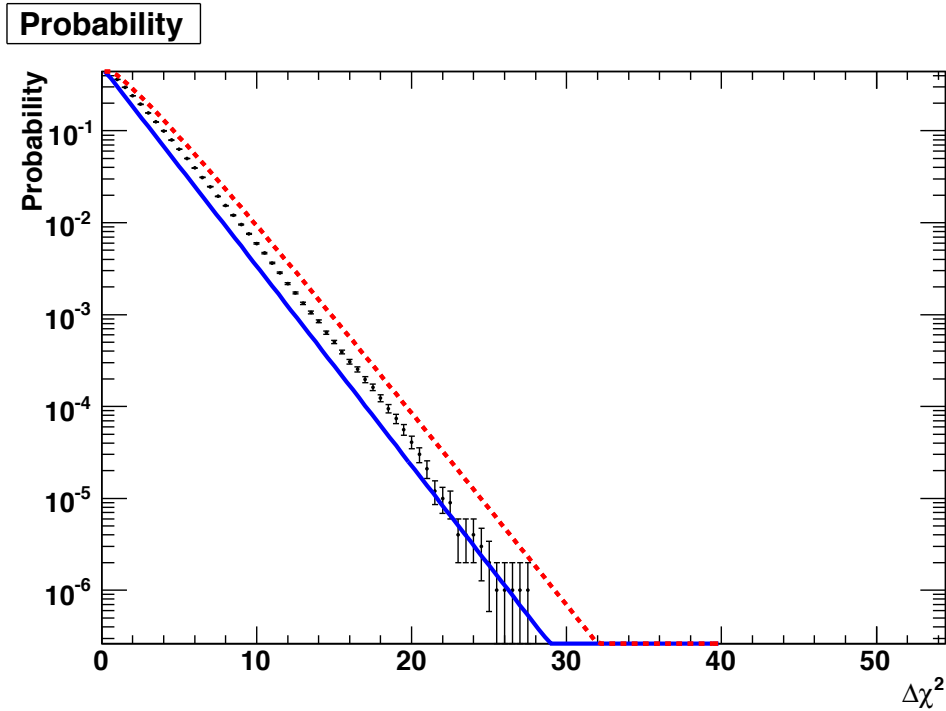
Figure 49: Probability to observe a background fluctuation with a given signal significance $\mathcal{S} \equiv \Delta\chi^2$ (plotted on the horizonal axis) anywhere in the range $6 < m_{A^0} < 7.5$ GeV.

- Υ *counting and luminosity*: provided by the Luminosity group. A preliminary estimate is 0.9% (TBC) [13] .

8.1 PDF Systematics

Systematics errors on the $e^+e^- \to \gamma\gamma$ PDF (fitted to the data sample without the IFR veto for the High-Energy region and to the MC sample of $e^+e^- \to \gamma\gamma$ events in the Low-Energy region) are statistical in nature. We take them into account by varying each parameter by its statistical uncertainty and observing the change in the fitted signal yield[6]. Correlations between parameters are taken into account.

Intrinsic uncertainties in the signal PDFs are small, but we need to take into account possible differences between the data and Monte Carlo. To that extent, we compare the PDF parameters between the data and MC for the $e^+e^- \to \gamma\gamma$ events in the High-Energy region near $m_X = 0$. The m_X^2 distribution is fitted to a Crystal Ball function, similar to the signal (the tail of the Crystal Ball is a bit longer for $e^+e^- \to \gamma\gamma$ events due to initial-state radiation). The comparison between the fits to the $\Upsilon(3S)$ sample and the SP-10 $\Upsilon(3S)$ MC is shown in Fig. 50, and the parameters are plotted in Fig. 51. Some of the variation (in particular, in μ) can be attributed to the differences in the simulated beam energy distributions (which is fixed to $m(\Upsilon(3S))$ in MC, but moves around in the data). We correct for the differences between the data and MC by applying corrections

---

[6]As implemented in `RooRarFit`
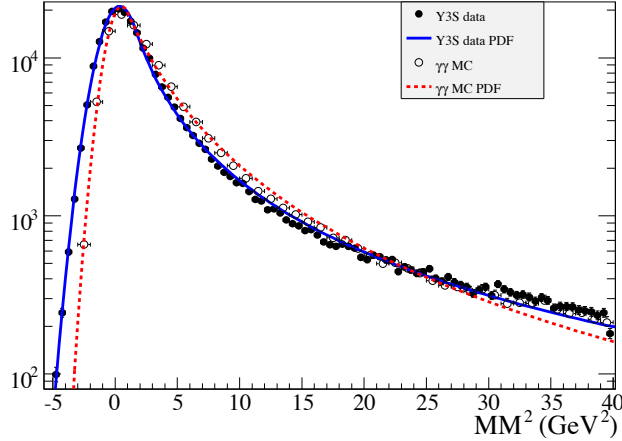
Figure 50: Overlay of the $e^+e^- \rightarrow \gamma\gamma$ PDF fitted to the $\Upsilon(3S)$ on-peak data (solid points and blue solid line) and SP10 MC (open circles and dashed red line).

Table 6: Corrections applied to the signal PDF

Parameter	Type	Correction
μ	Additive	-0.153 ± 0.077
σ	Multiplicative	1.34 ± 0.17
α	Additive	-0.36 ± 0.18
n	Multiplicative	0.72 ± 0.14

to all 4 Crystal Ball parameters for the signal PDF: additive corrections for parameters μ and α, and multiplicative correction to σ and n parameters. The corrections are based on the differences between the full on-peak $\Upsilon(3S)$ dataset and SP-10 $e^+e^- \rightarrow \gamma\gamma$ MC (produced on $\Upsilon(3S)$), and are listed in Table 6. We assign half of the difference as a systematic error for a parameter, ignoring correlations (this is conservative).

8.2 Trigger and Filter Systematics

We evaluate the efficiency of the L1/L3 trigger and BGFilters using a sample of prescaled trigger passthru events (`DigiFL1Open||DigiFL3Open`). We measure the efficiency of the combinations `DigiFSingleGamma&&BGFilterSingleGammaInvisible` (high-energy region) and `DigiFSingleGammaLowE&&BGFilterSingleGammaInvisibleLowE` (low-energy region) on a sample of events in the pass-thru dataset that pass the standard analysis selection (without the IFR veto for the high-energy region). The results are:

- High-Energy region: $\varepsilon_{\text{trig}} = 0.9985 \pm 0.0011$ (data), $\varepsilon_{\text{trig}} = 0.99971 \pm 0.00008$ ($e^+e^- \rightarrow \gamma\gamma$ MC). We ignore the difference and assign a systematic uncertainty of 0.1%.
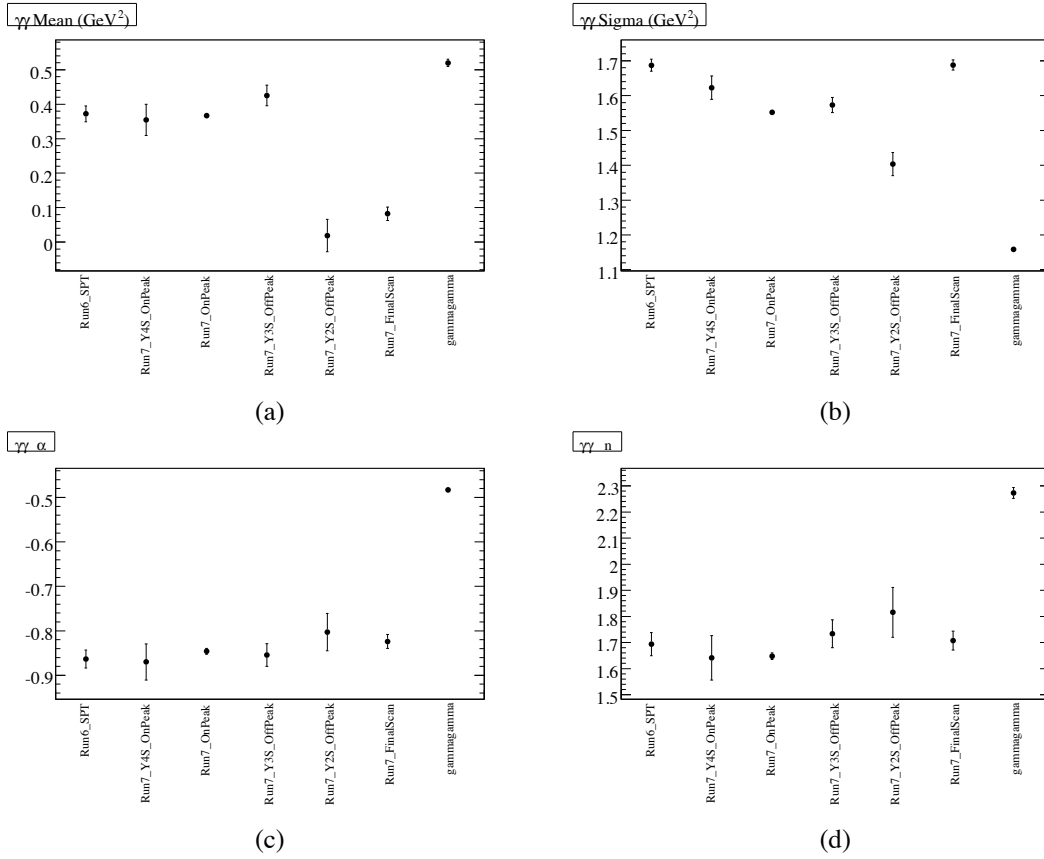
Figure 51: Comparison between parameters of the $e^+e^- \to \gamma\gamma$ PDF, fitted to different data samples and SP-10 MC. (a) μ (mean of the Gaussian part of the Crystal Ball function), (b) σ (Gaussian resolution), (c) α (transition point) and (d) n (power law).

- Low-Energy region: $\varepsilon_{\rm trig} = 0.9775 \pm 0.0044$ (data), $\varepsilon_{\rm trig} = 0.9824 \pm 0.0013$ (radiative bhabha MC). We ignore the difference and assign a systematic uncertainty of 0.4%.

9 Fits to the Full $\Upsilon(3S)$ Dataset

9.1 High-Energy Region

We use the full dataset `AllEvents-Run7-R24-Y3S-OnPeak-v04`, which contains $28.6\,{\rm fb}^{-1}$, or $(121.9 \pm 1.1) \times 10^6$ $\Upsilon(3S)$ decays. The high-energy/low-mass region $(-5 < m_X^2 < 40\,{\rm GeV}^2)$ contains 955 events (with the IFR veto cut) and $N_{\gamma\gamma\,{\rm EMC}}^{\rm on} = 244462$ without the IFR veto. The background yields scale reasonably well with luminosity (the ratios between the full and ''LowHigh'' datasets are 10.19 ± 0.07 and 11.6 ± 1.3 for the samples without and with the IFR cut, respectively, compared to the luminosity ratio of 11.6). We fix the $e^+e^- \to \gamma\gamma$ contribution to

$$N_{\gamma\gamma} = 244462 \times (4.5 \pm 1.9) \times 10^{-4} = 110 \pm 46 \tag{6}$$

The yields of the signal and continuum background events are floated, as well as the shape parameter c of the continuum background

$$f_{\rm bkg}(m_X^2) \propto \exp(cm_X^2) \tag{7}$$

We first fit the entire region to the combination of $e^+e^- \to \gamma\gamma$ and the continuum background PDFs (*i.e.* fix signal contribution at zero). This tests whether the gross features of the background PDF are fitted well. We float the $e^+e^- \to \gamma\gamma$ efficiency, and the fit returns $\varepsilon_{\rm IFR} = 3.5^{+7}_{-6} \times 10^{-4}$, consistent with the estimates from the off-resonance data. The fit is shown in Fig. 52. We see that the fit describes the data well, including the region near $m_X^2 = 0$. No points are more than 2σ away from the fit, and the overall $\chi^2/df = 26.5/42$.

The only suspect region of the background-only fit just above the $m_X^2 = 0$ peak, where several points in a row are above the curve. This could be a fluctuation, or a hint of the ISR production of hadronic states above $m_X = 1$ GeV. To test that hypothesis, we perform a fit with a floated yield, mean, and the tail parameter n of the $e^+e^- \to \gamma\gamma$ PDF. We find the shift in the mean position of the nominal $e^+e^- \to \gamma\gamma$ peak of $\Delta\mu = +0.36 \pm 0.39$, and a shift in the Crystal Ball tail parameter $\Delta n = -0.1^{+0.0}_{-1.1}$ (the MINOS step terminated without finding the positive error). We conclude that there is no strong evidence for the additional ISR component, but conservatively increase the uncertainty on the mean and the tail parameters of the $e^+e^- \to \gamma\gamma$ PDF to $\sigma(\mu) = 0.39$ and $\sigma(n) = 0.1$.

Fig. 55 shows the results of the fits at different values of m_{A^0}. We show the statistical error and the total uncertainty in the yield, taking into account the systematic errors. There are no yields above 3σ. The fitted background shape parameters are shown in Fig. 54.

A fit with a floated m_{A^0} converges on $m_{A^0} = 5.22 \pm 0.06$ GeV and $N_{\rm sig} = 37 \pm 15$, with the statistical significance of 2.6σ. This fit is shown in Fig. 53.

Since no significant signal is found in this region, we compute the upper limits on the branching ratio of $\Upsilon(3S) \to \gamma A^0$, . The 90% C.L. Bayesian upper limits (computed with a uniform prior and a gaussian likelihood function) for the full $\Upsilon(3S)$ dataset are shown in Fig. 56.
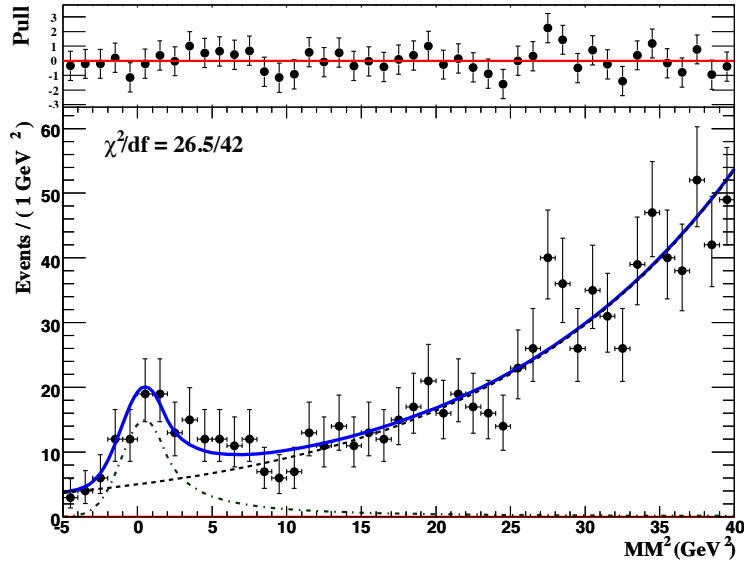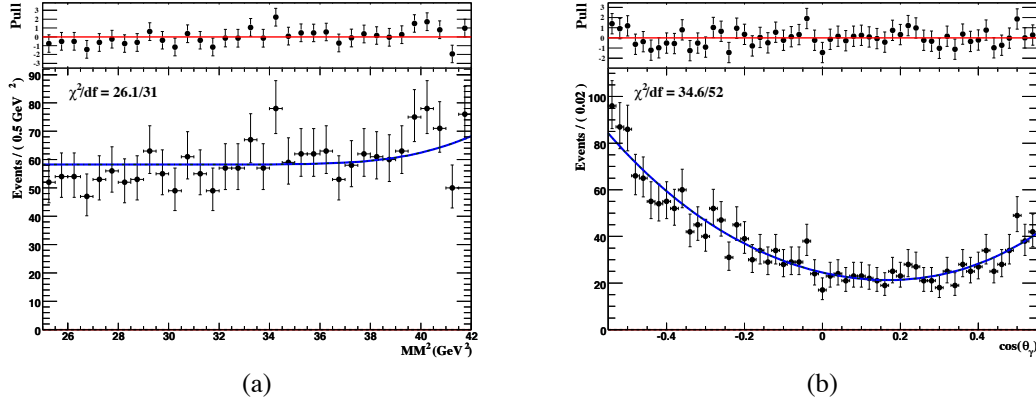
Figure 52: Fit the the full $\Upsilon(3S)$ dataset with signal yield fixed to zero. The bottom plot shows the data (solid points) overlaid by the full PDF curve (solid blue line), $e^+e^- \to \gamma\gamma$ contribution (dot-dashed green line), and continuum background PDF (black dashed line). The top plot shows the pulls.
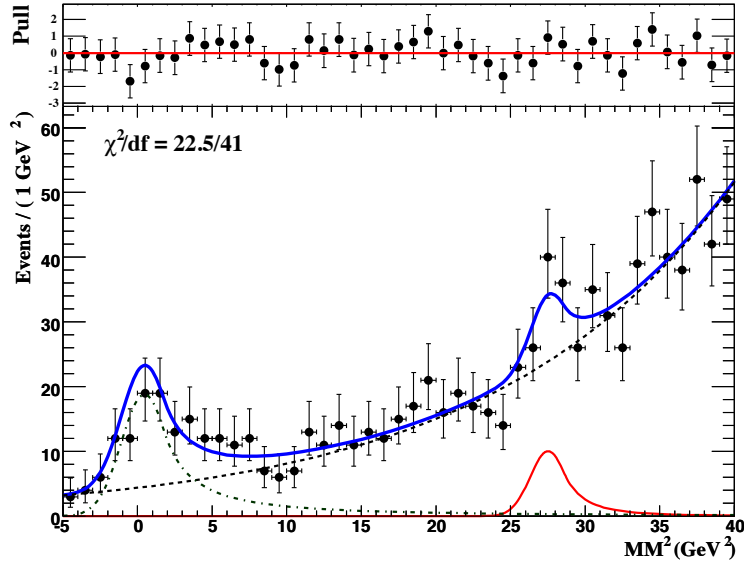
Figure 53: Fit the the full $\Upsilon(3S)$ dataset with floated m_{A^0}. The bottom plot shows the data (solid points) overlaid by the full PDF curve (solid blue line), signal contribution with $m_{A^0} = 5.23 \pm 0.06$ GeV (solid red line), $e^+e^- \to \gamma\gamma$ contribution (dot-dashed green line), and continuum background PDF (black dashed line). The top plot shows the pulls.

48
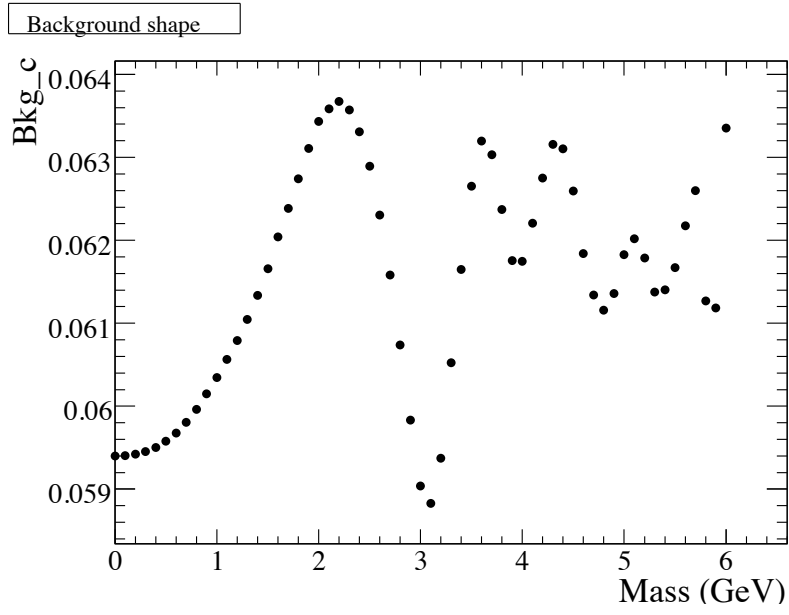
Figure 54: Exponential background parameter `Bkg_c` as a function of assumed mass m_{A^0} for the full $\Upsilon(3S)$ sample. Statistical uncertainties are shown (but are smaller than the size of the point).
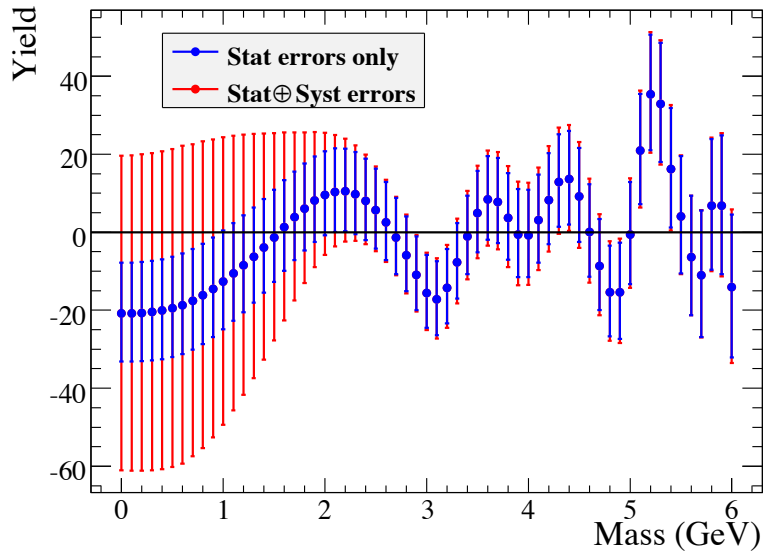
Figure 55: Signal yields N_{sig} as a function of assumed mass m_{A^0} for the full $\Upsilon(3S)$ sample. Blue error bars are statistical only, and the red error bars include the systematic contributions.
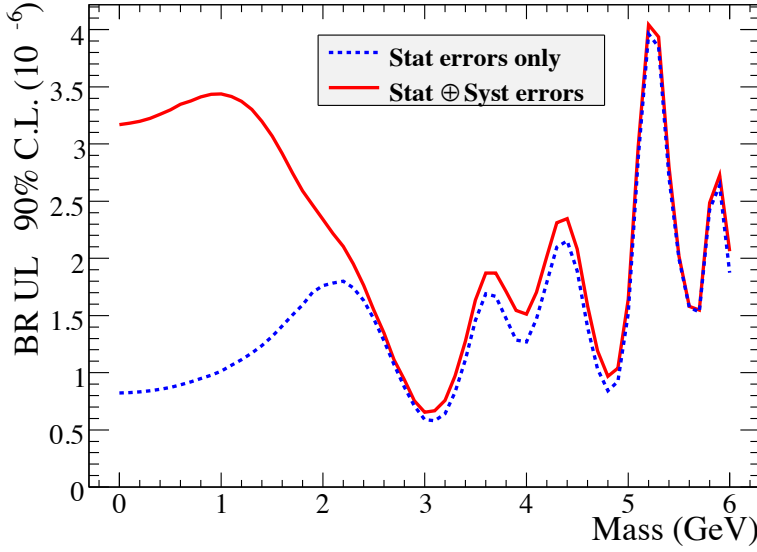
Figure 56: 90% C.L. upper limits for the full Run7 $\Upsilon(3S)$ data sample. The dashed blue line shows the statistical uncertainties only, the solid red line includes the systematic uncertainties.

9.2 Low-Energy Region: Pre-Unblinding

We use the dataset `BGFilterSkim-Run7-Y3S_OnPeak-R24-v01`, which contains $19.6\,\mathrm{fb}^{-1}$, or $(83.0 \pm 0.7) \times 10^6\,\Upsilon(3S)$ decays. The low-energy/high-mass region ($30 < m_X^2 < 70\,\mathrm{GeV}^2$) contains 236743 events with $|\cos(\theta_\gamma^*)| < 0.6$ and 177866 events with $|\cos(\theta_\gamma^*)| < 0.55$. The ratio of the yields between the full `BGFilterSkim` and `BGFilterSkim-High` is 8.33 ± 0.05, compared to the luminosity ratio of 8.17. We fit the low-energy dataset as described in Section 6, but we first keep the region around $m_{A^0} = 6.7\,\mathrm{GeV}$ blind.

We first fit the region $25 < m_X^2 < 42\,\mathrm{GeV}^2$ (extended a little below the nominal cutoff of $30\,\mathrm{GeV}^2$ to increase statistics) to the combination of $e^+e^- \to \gamma\gamma$ and the continuum background PDFs (*i.e.* fix signal contribution at zero). The fit is shown in Fig. 57. There are no features in the data, and the fit converges well. No points are more than 3σ away from the fit, and the overall χ^2 is good. A fit with a floated m_X^2 converges on $m_{A^0} = 5.04 \pm 0.12\,\mathrm{GeV}$, $N_{\mathrm{sig}} = 23 \pm 37$, and significance of 0.6σ.

We also fit the region $50 < m_X^2 < 70\,\mathrm{GeV}^2$ to the combination of $e^+e^- \to \gamma\gamma$ and the continuum background PDFs (*i.e.* fix signal contribution at zero). Here the full-statistics fit reveals two potential pathologies. First, since we had to place the lower cut at $50 < m_X^2\,\mathrm{GeV}^2$ to blind the previously observed peak, the fitter looses any sensitivity to the shape of the $e^+e^- \to \gamma\gamma$ background, or to the peaking structure of the bhabha background. Therefore, we fix the parameters of the $e^+e^- \to \gamma\gamma$ PDF to the Monte Carlo values, and fix the bhabha peak magnitude to zero.

A second problem is that when the data are fit over a broad range $50 < m_X^2 < 70\,\mathrm{GeV}^2$ where the bhabha rate changes by several orders of magnitude, it becomes apparent that the single-
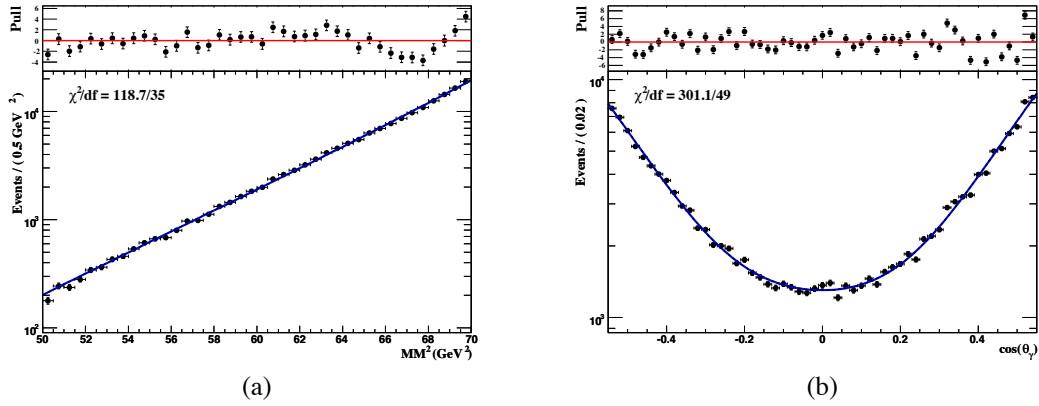
(a)　　　　　　　　　　　　　　(b)

Figure 57: Fit the the full `BGFilterSkim` dataset with signal yield fixed to zero. The bottom plot shows the data (solid points) overlaid by the full PDF curve (solid blue line), $e^+e^- \to \gamma\gamma$ contribution (dot-dashed green line), and radiative bhabha background PDF (black dashed line). The top plot shows the pulls.

exponent shape of the bhabha background is inadequate. We therefore describe the bhabha PDF in m_X^2 as an exponentiated polynomial:

$$f_{\text{bhabha}}(m_X^2) \propto \exp(c_1 m_X^2 + c_2 m_X^4 + c_3 m_X^6) \tag{8}$$

which parameters $c_{[1..3]}$ floated. When performing fits over a more narrow range (*e.g.*, when scanning for the yield as a function of m_{A^0}), it may be appropriate to fix some of the c_i coefficients to zero.

The fit is shown in Fig. 58. The fit has 9 free parameters: 3 coefficients c_i for $f_{\text{bhabha}}(m_X^2)$, 4 polynomial coefficients for $f_{\text{bhabha}}(\cos\theta_\gamma^*)$, N_{bhabha}, and $N_{\gamma\gamma}$. The fit converges well (albeit slowly) with the full error matrix. There is a feature in the plot of the residual pulls around $m_X^2 = 64 \text{ GeV}^2$, which worsens the overall χ^2. This corresponds to a photon energy of $E_\gamma^* \approx 2.1$ GeV, and could be due to an onset of `L3OutSingleGamma` line. Signal fits with m_{A^0} fixed or floated around that region consistently return negative yields, so the structure is not likely due to any signal-like components. We will likely need to add a threshold-like function to the m_X^2 PDF for the background to accommodate this feature.

The behavior of the fit at lower masses seems under control. However, the plot of the polar angle $\cos\theta_\gamma^*$ exhibits poor χ^2, due to the effects of the EMC crystal edges. How these binning effects influence the unbinned likelihood fit also needs to be understood.

9.3 Low-Energy Region: Full Unblinding

A close look at the distributions at high values of m_X^2 reveals a strange feature in the `DigiFSingleGamma` line (same as L3 line `L3OutSingleGamma`). As shown in Fig. 59a, the events pass `L3OutSingleGamma` even though the photon candidates are *below* the nominal trigger threshold of 2 GeV. Fig. 59b shows that these events are likely due to hot towers not masked in L3. In any case, understanding the effects of these events will take some time.
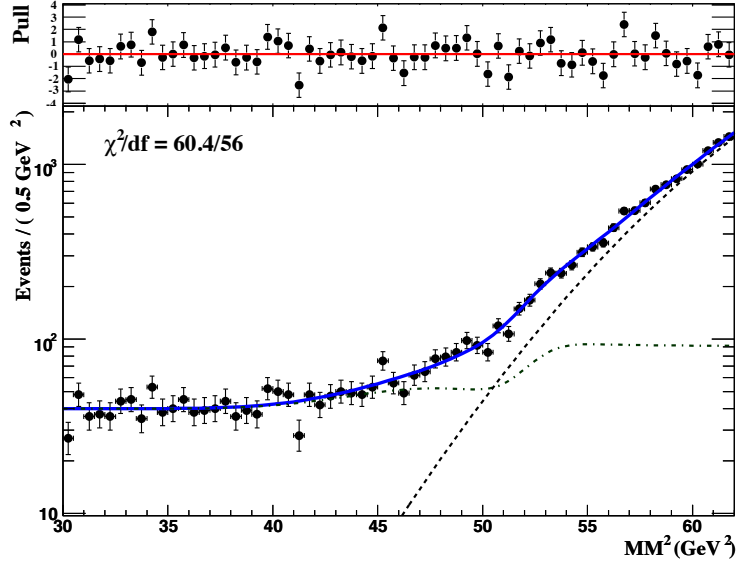
51

(a)   (b)

Figure 58: Fit the the the full `BGFilterSkim` dataset with signal yield fixed to zero. The bottom plot shows the data (solid points) overlaid by the full PDF curve (solid blue line), $e^+e^- \to \gamma\gamma$ contribution (dot-dashed green line), and radiative bhabha background PDF (black dashed line). The top plot shows the pulls.


(a)   (b)
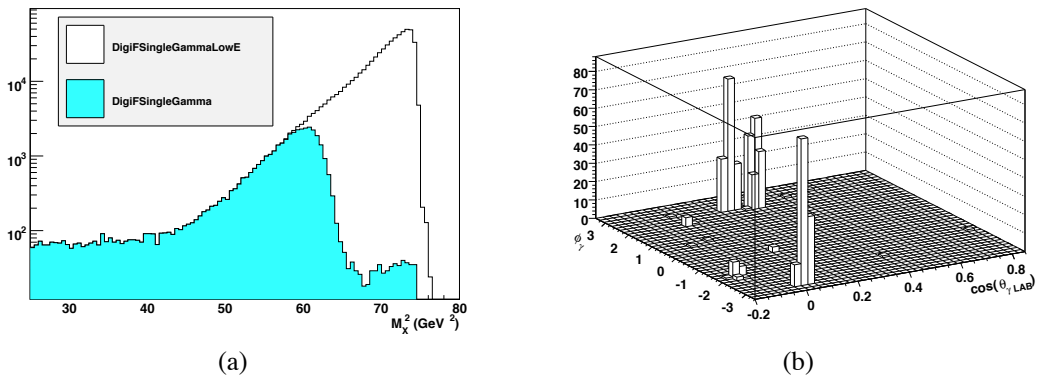
Figure 59: (a): distribution of m_X^2 for events passing L3 line `L3OutSingleGamma` (cyan histogram) and events passing `L3OutSingleGammaLowE` (open histogram). The tail at high values of m_X^2 is likely due to the hot L3 towers, as shown in (b): the distribution $(\cos\theta_\gamma, \phi)$ of the photon candidates in events with $m_X^2 > 68 \ \mathrm{GeV}^2$ passing `L3OutSingleGammaLowE`.
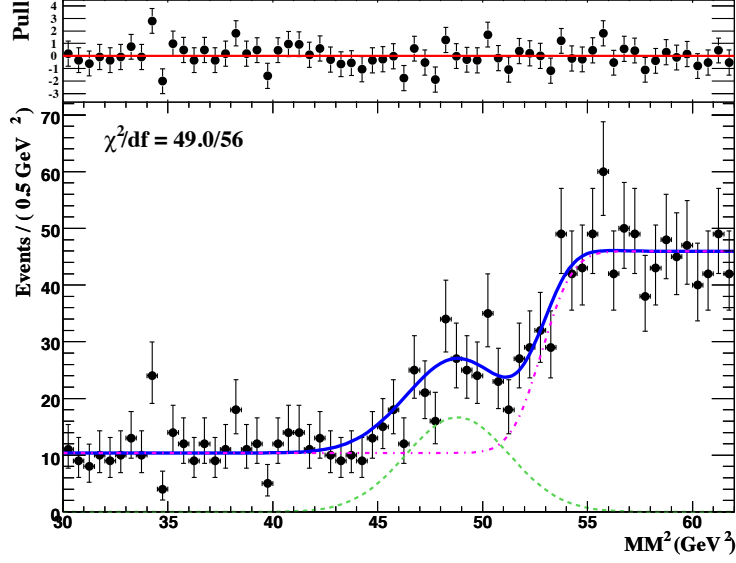
Figure 60: Projection of the 2d fit to the Monte Carlo $e^+e^- \to \gamma\gamma$ events onto the missing mass squared m_X^2. The bottom plot shows the data (solid points) overlaid by the full PDF curve (solid blue line), the peaking contribution (dashed green line), and the continuum component (pink dot-dashed line). The top plot shows the pulls.

Likewise, correcting for the effects of EMC crystal edges on the 2-dimensional likelihood fit will require either a substantial toy Monte Carlo study, or a data calibration sample (the latter is being developed). Therefore, we perform the preliminary analysis with in the low-energy region with a *1-dimensional* fit to the m_X^2 distribution over a *restricted range* $30 \leq m_X^2 \leq 62$ GeV^2 and $|\cos\theta_\gamma^*| < 0.46$. The cut on $\cos\theta_\gamma^*$ was picked to maximize $S/\sqrt{B}$ figure-of-merit, using the signal angular distribution for S, and the angular distribution of the background events B fitted to the off-$\Upsilon(3S)$ sample. Signal efficiency for this selection is $\varepsilon = 20.65\%$.

The full `BGFilterSkim` data sample contains 16806 in the fit region. We perform the fit for each value of assumed m_{A^0} over the entire range $30 \leq m_X^2 \leq 62$ GeV^2 to avoid edge effects. To allow maximum flexibility in the fit while maintaining reliable convergence, we parameterize the radiative bhabha distribution as an exponentiated 2nd-order polynomial:

$$f_{\text{bhabha}}(m_X^2) \propto \exp(c_1 m_X^2 + c_2 m_X^4) \tag{9}$$

We also carefully determine the PDF for the $e^+e^- \to \gamma\gamma$ component with a high-statistics ($\approx 1\times$ data luminosity) Monte Carlo sample. The $e^+e^- \to \gamma\gamma$ PDF is parameterized as a sum of a broad Gaussian (peaking component from forward conversions) and a flat continuum with a threshold starting around $m_X^2 = 53$ GeV^2 (showers in the central region). In order to disentangle these contributions better, we fit the m_X^2 and $\cos\theta_\gamma^*$ distributions simultaneously, where the peaking component is parameterized as a highly-peaked 4th-order polynomial, and the continuum component has $1 + p_2 \cos^2\theta_\gamma^*$ distribution. The fit to the $e^+e^- \to \gamma\gamma$ Monte Carlo sample is shown in Fig. 60. We then keep m_X^2 component of the fit.

In the data fits, we fix four parameters of the $e^+e^- \to \gamma\gamma$ PDF (location and width of the

threshold in the continuum part, and location and width of the peak component) to the values determined in the Monte Carlo fit (Fig. 60), varying them by $\pm 1\sigma$ to determine the systematic uncertainties. Two other parameters of the $e^+e^- \to \gamma\gamma$ PDF (magnitude of the threshold $\mathcal{A}_{\text{threshold}}$ in the continuum part and the fraction of the peaking component f_{peak}) are floated. The parameters of the bhabha PDF are floated, as are the yields of the $e^+e^- \to \gamma\gamma$ and bhabha contributions. The fit to the entire sample, with the signal yield fixed at zero, is shown in Fig. 61.

The results of the fits with the floated signal yield, in fine steps of m_{A^0}, are shown in Fig. 63. The floated background parameters are shown in Fig. 62.

We do not observe significant access in the full dataset. The point at $m_{A^0} = 6.7$ GeV yields $N_{\text{sig}} = 38 \pm 40$. It appears the access seen in the BGFilterSkim-High dataset was a fluctuation. We compute the upper limits on the branching ratio of $\Upsilon(3S) \to \gamma A^0$, . The 90% C.L. Bayesian upper limits (computed with a uniform prior and a gaussian likelihood function) as a function of m_{A^0} are shown in Fig. 64. The sensitivity just above the threshold region in the $e^+e^- \to \gamma\gamma$ PDF ($m_X \approx 57$ GeV^2) deteriorates, as the signal and $e^+e^- \to \gamma\gamma$ yields become strongly correlated (this is where a 2d fit would have helped). Nonetheless, we set the limits from 2×10^{-6} to 3×10^{-5} for the entire region of interest, or about an order of magnitude better than the previous results from CLEO[9].

Figure 61: 1d fit to the the the full BGFilterSkim dataset with signal yield fixed to zero. The bottom plot shows the data (solid points) overlaid by the full PDF curve (solid blue line), $e^+e^- \to \gamma\gamma$ contribution (dot-dashed green line), and radiative bhabha background PDF (black dashed line). The top plot shows the pulls.
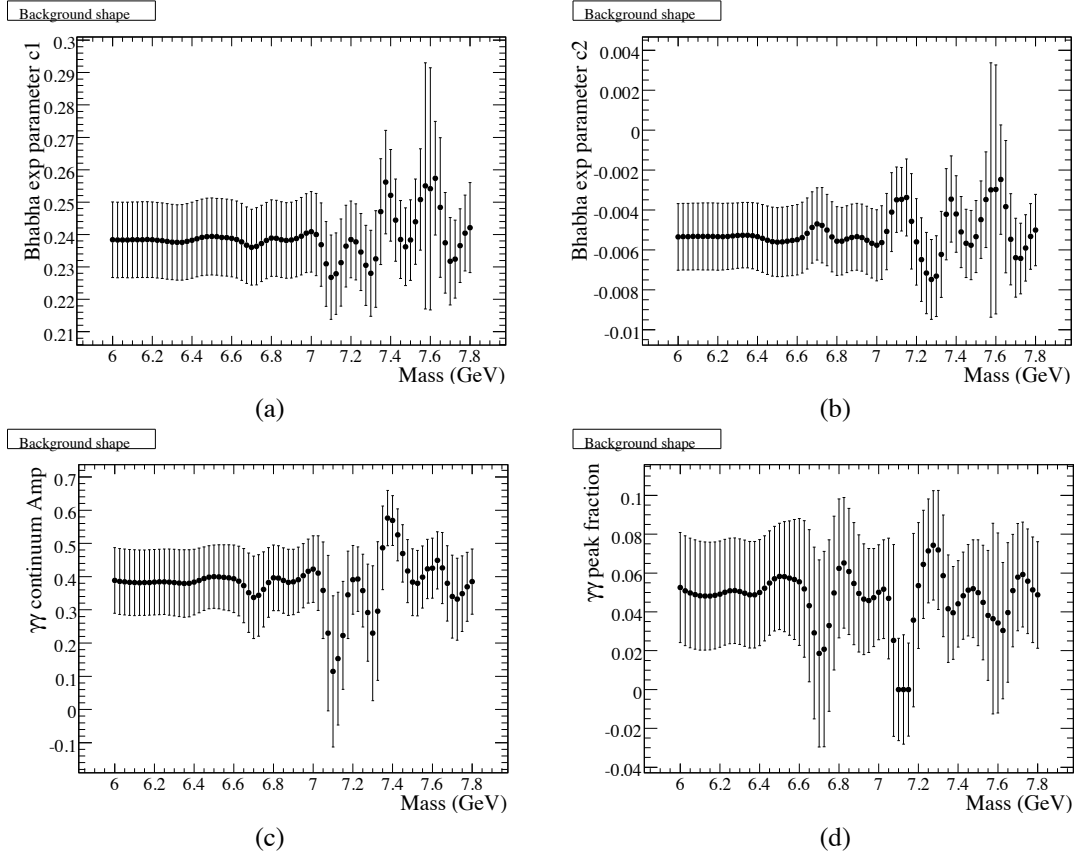
Figure 62: Parameters of the bhabha PDF c_1 (a), c_2 (b) and of the $e^+e^- \to \gamma\gamma$ PDF $\mathcal{A}_{\text{threshold}}$ (c) and f_{peak} as a function of assumed mass m_{A^0} for the full $\Upsilon(3S)$ sample. Statistical uncertainties are shown.
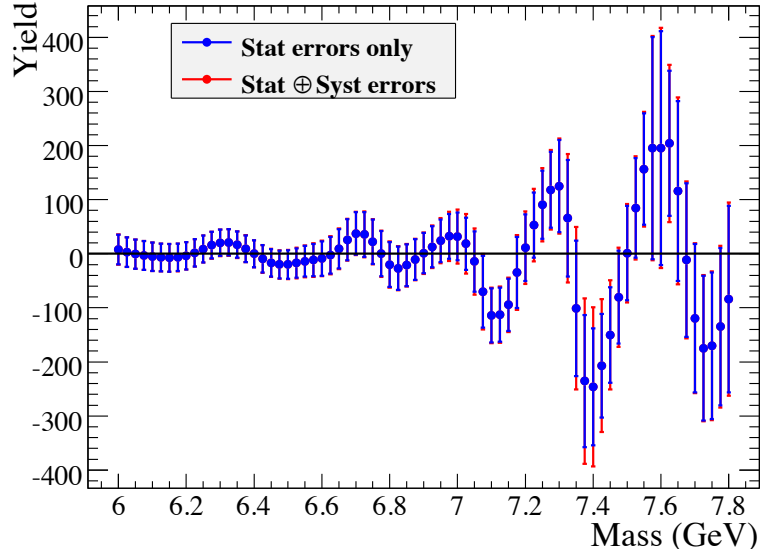
Figure 63: Signal yields N_{sig} as a function of assumed mass m_{A^0} for the full $\Upsilon(3S)$ sample. Blue error bars are statistical only, and the red error bars include the systematic contributions.
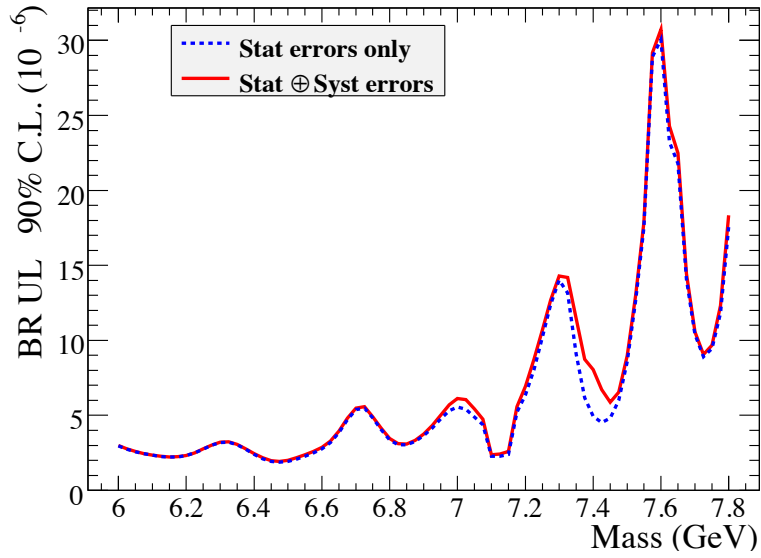
Figure 64: 90% C.L. upper limits for the full Run7 $\Upsilon(3S)$ data sample in the mass range $6 < m_{A^0} < 7.8$ GeV. The dashed blue line shows the statistical uncertainties only, the solid red line includes the systematic uncertainties.

10 Acknowledgments

References

[1] C. Patrignani *et al*., *The Physics Case for Running the B-factory at the $\Upsilon(3S)$ Resonance*, BAD1928 (2008).

[2] R. Barate *et al*., LEP Working Group for Higgs boson searches, Phys. Lett. **B565**, 61 (2003).

[3] EWWG, Phys. Rept. **427**, 257 (2006).

[4] R. Dermisek and J.F. Gunion, Phys. Rev. Lett. **95**, 041801 (2005).

[5] F. Wilczek, Phys. Rev. Lett. **39**, 1304 (1977).

[6] R. Dermisek, J.F. Gunion, and B. McElrath, Phys. Rev. D **76**, 051105 (2007).

[7] J.F. Gunion, D. Hooper, and B. McElrath, Phys. Rev. D **73**, 015011 (2006).

[8] H. Park *et al*., HyperCP Collaboration, Phys. Rev. Lett. **94**, 021801 (2005).

[9] R. Balest *et al*., Phys. Rev. D **51**, 2053 (1995).

[10] W.-M. Yao *et al*., *Review of Particle Physics*, J. Phys. G **33**, 1 (2006).

[11] A. Down, *The Search for Light Dark Matter at the Stanford Linear Accelerator*, Honors Thesis, University of California, Berkeley (2007).

[12] Frank Porter, private communication.

[13] G. McGregor and C. Hearty, BAD 2069.

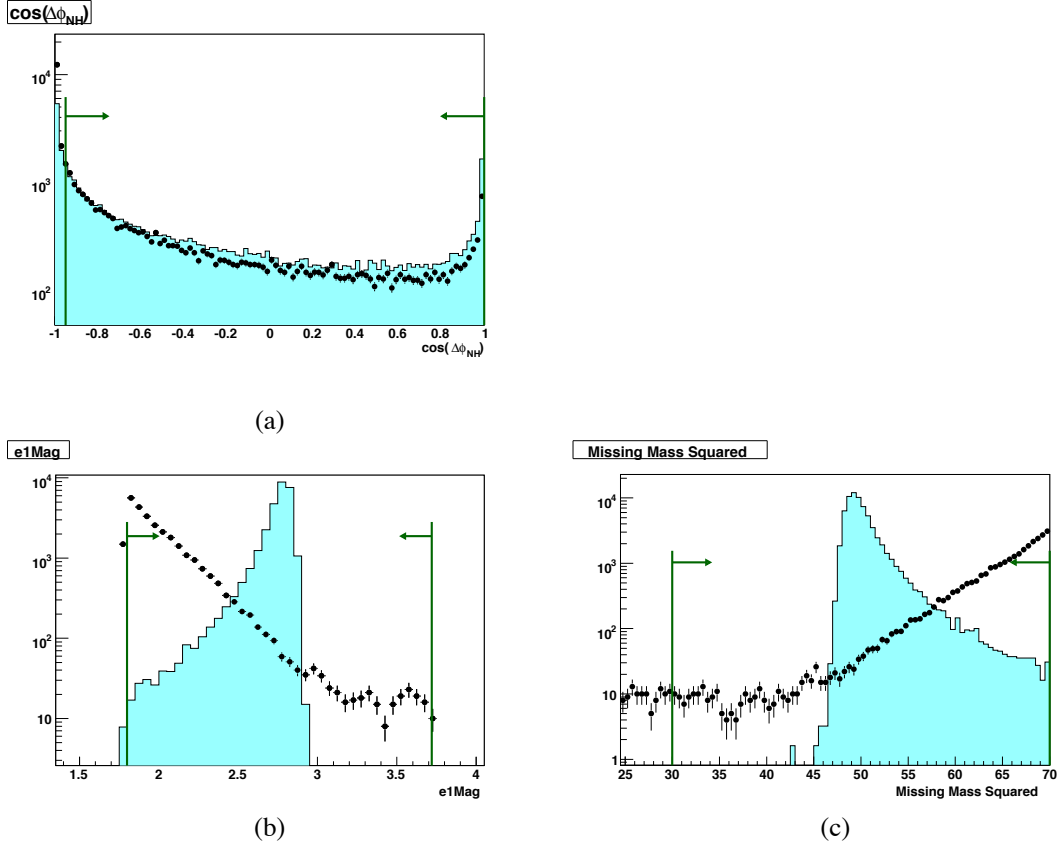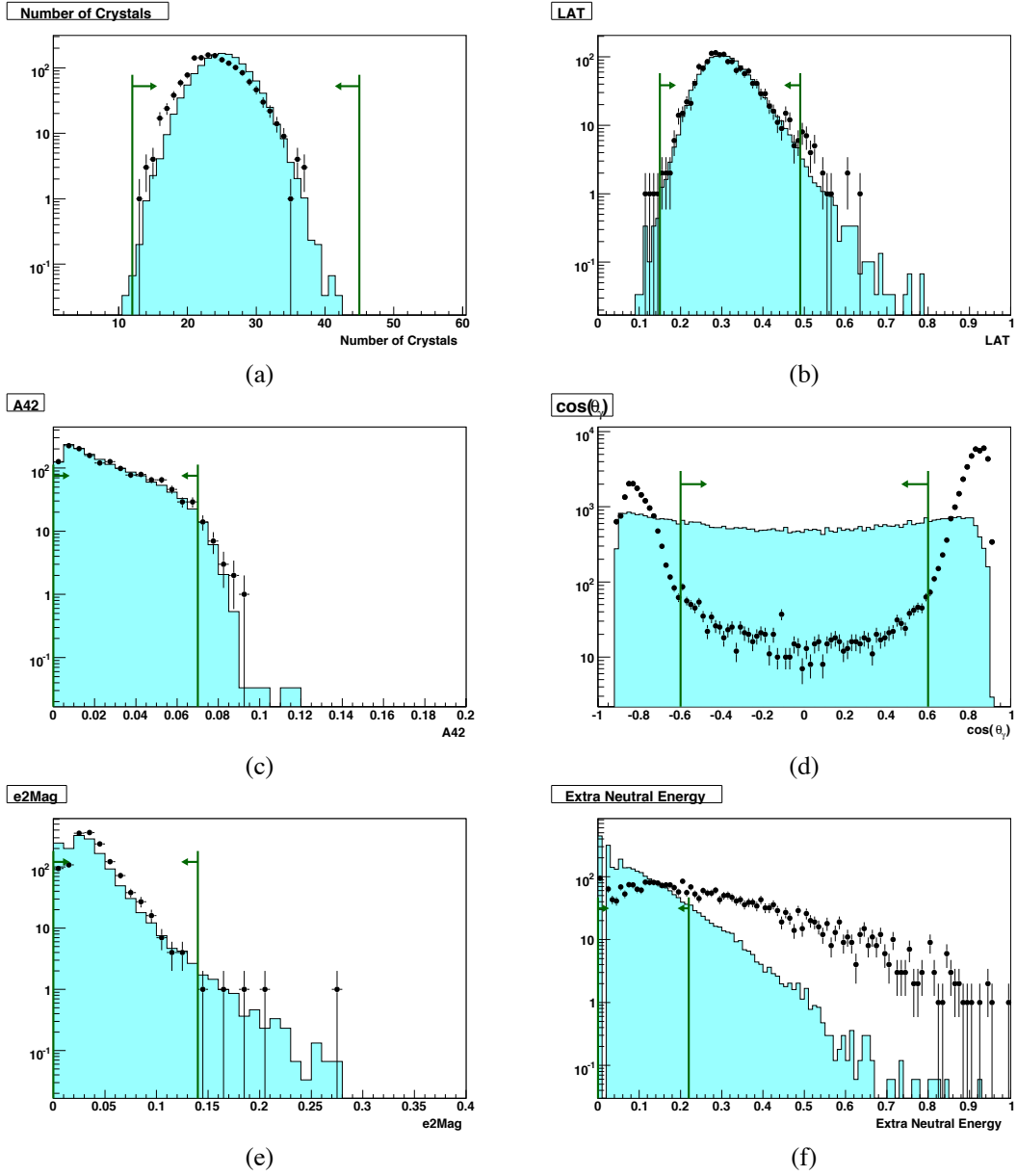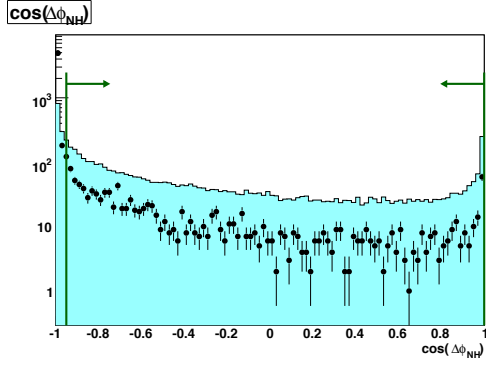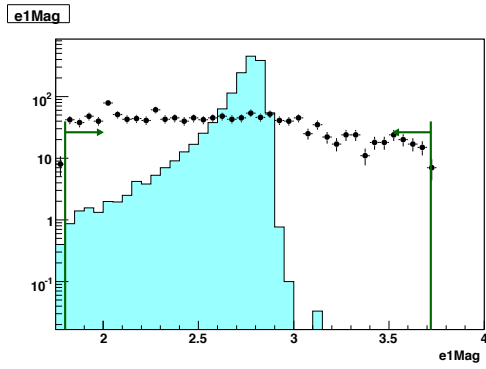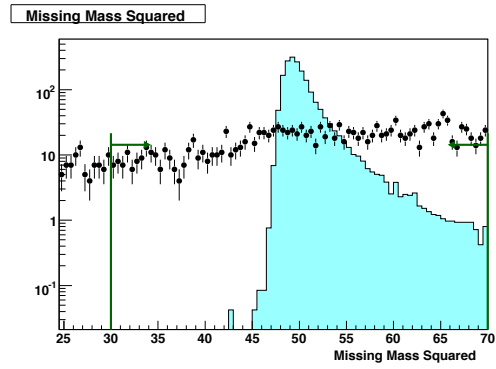# A $N - 1$ **Selection Plots**
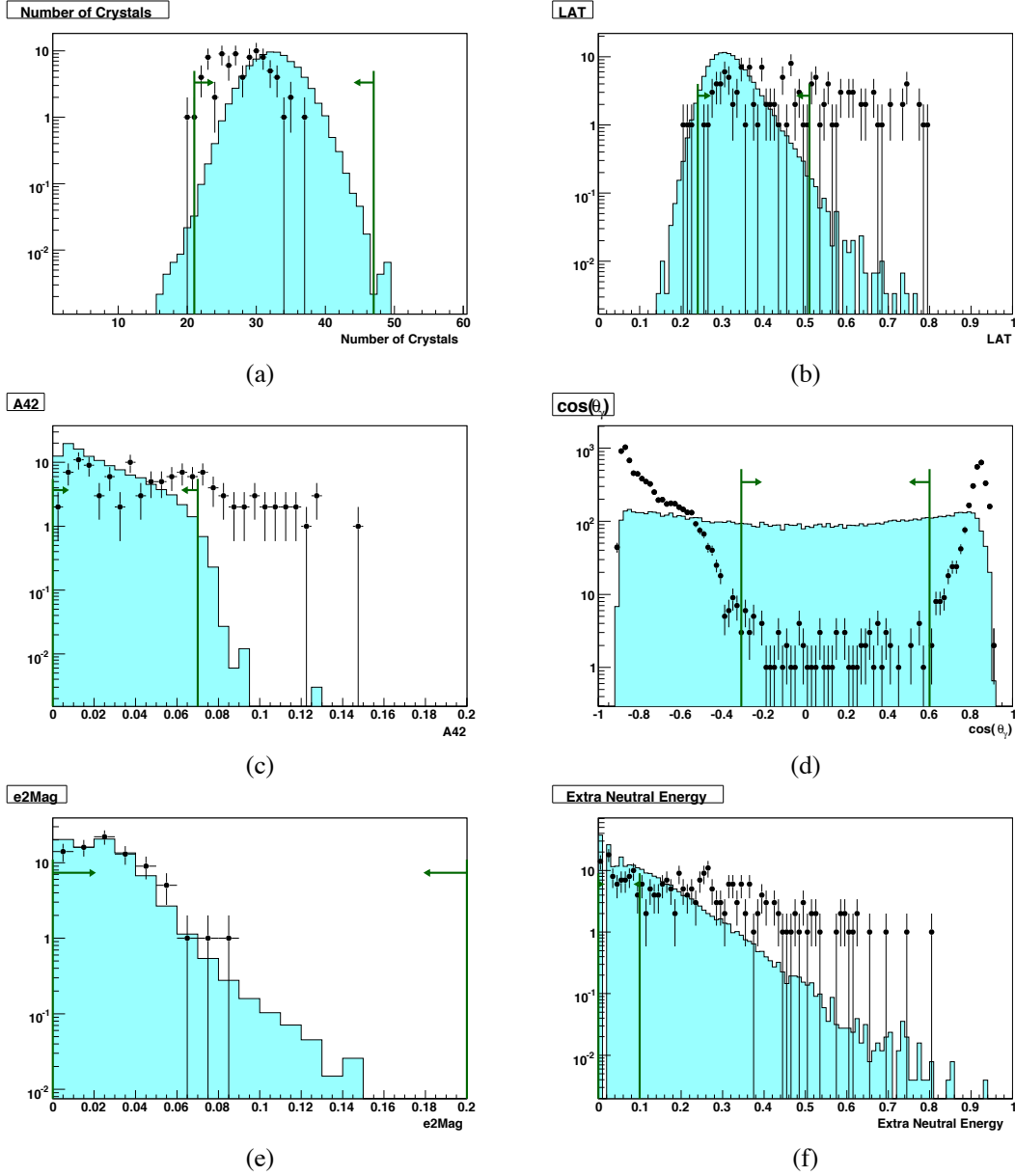
Figure 65: Plots of each of the selection variables with all final cuts, except on the variable plotted, applied. The selections for the high energy region, $3.2 < E_\gamma^* < 5.5$ are applied. The points represent the $2.5\ \text{fb}^{-1}$ "HighLow" dataset, and the cyan histogram corresponds to the $m_{A^0} = 2\ \text{GeV}$ Monte Carlo sample SP-8740. (a): number of crystals in the EmcBump; (b): LAT moment; (c): a_{42} moment; (d): $\cos\theta_\gamma^*$; (e): E_2; (f) E_{extra}.
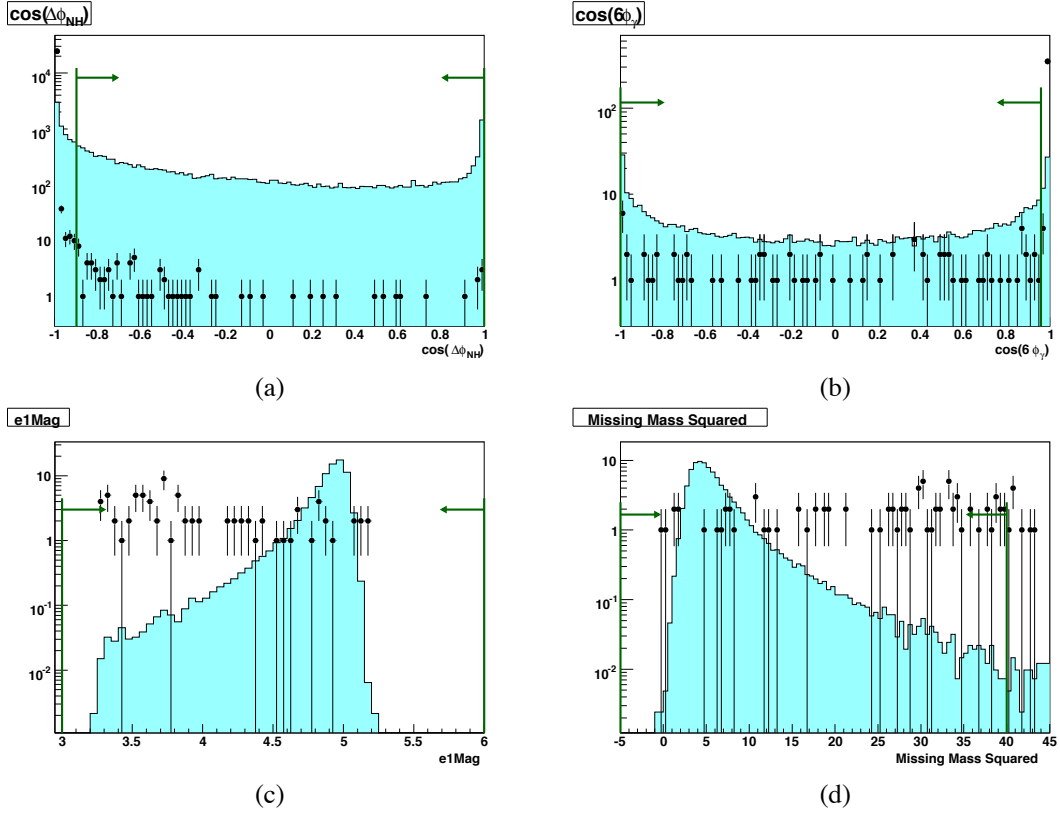
Figure 66: Plots of each of the selection variables with all final cuts, except on the variable plotted, applied. The selections for the high energy region, $3.2 < E^*_\gamma < 5.5$ are applied. The points represent the 2.5 fb^{-1} "HighLow" dataset, and the cyan histogram corresponds to the $m_{A^0} = 2$ GeV Monte Carlo sample SP-8740. The green lines show the selection interval. (a): $\cos \Delta\phi_{NH}$; (b): $\cos(6\phi_\gamma)$; (c): $< E^*_\gamma$; (d): recoil mass squared m^2_X.
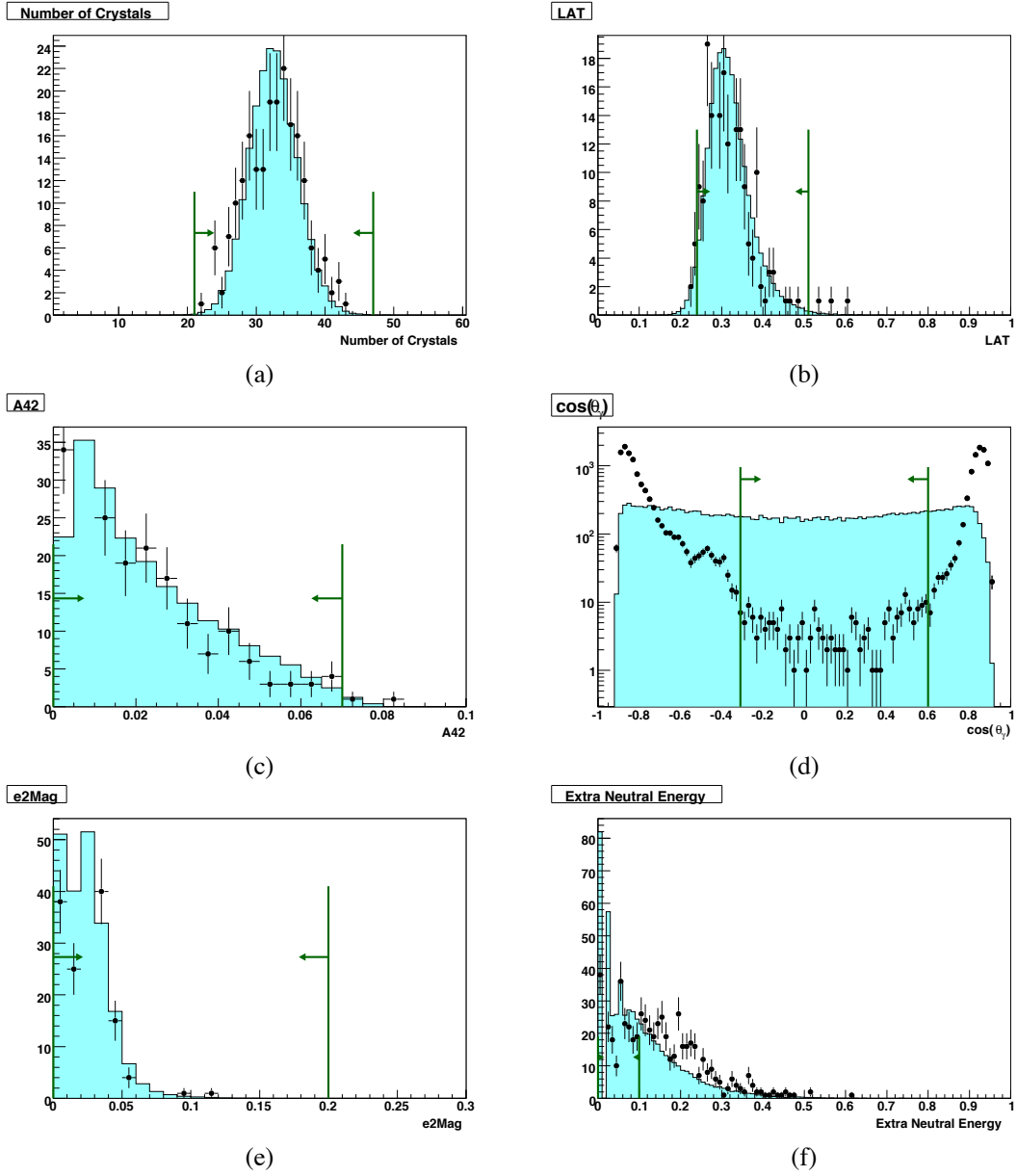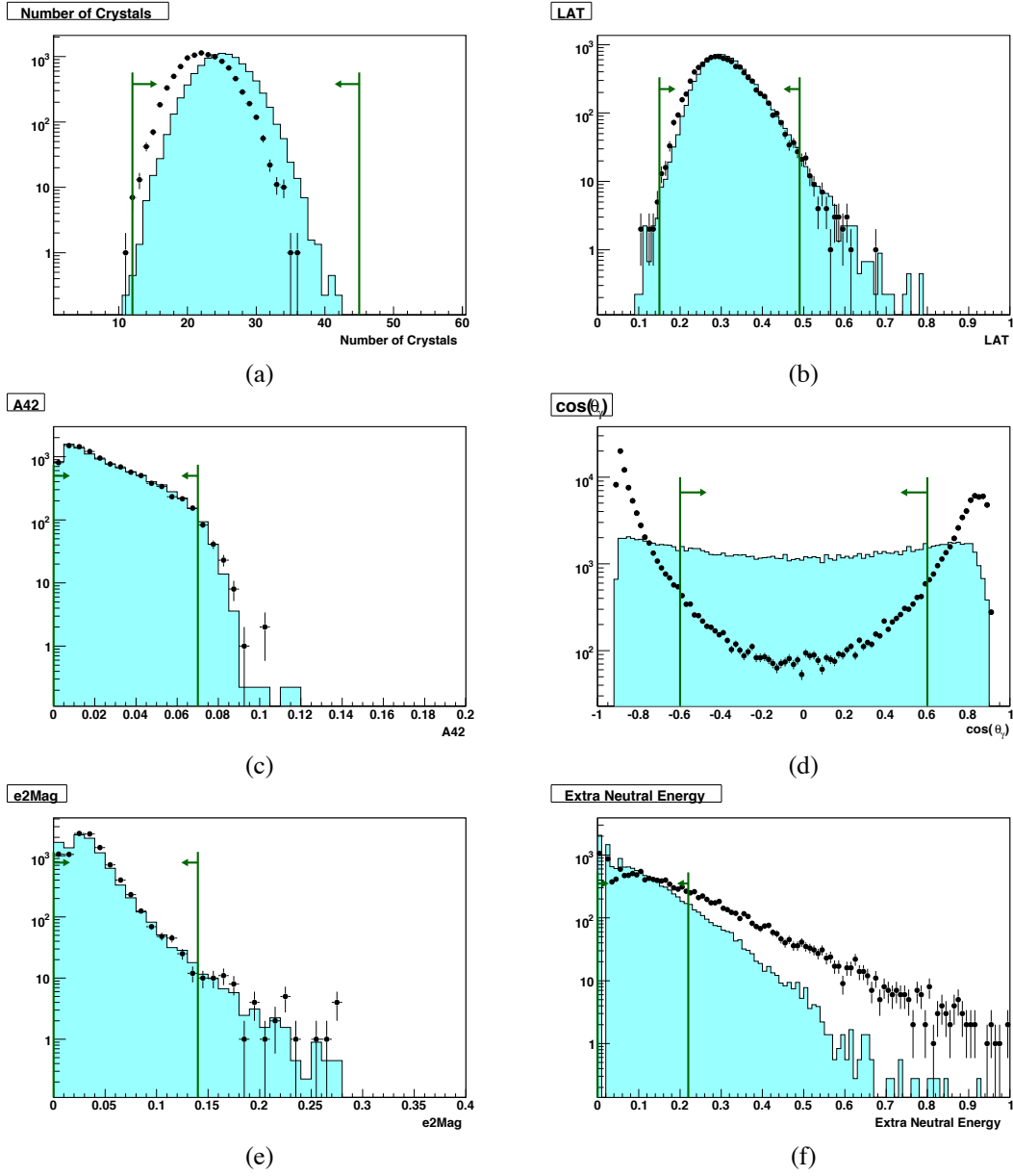
Figure 67: Plots of each of the selection variables with all final cuts, except on the variable plotted, applied. The selections for the high energy region, $3.2 < E_\gamma^* < 5.5$ are applied. The points represent the $e^+e^- \to \gamma\gamma$ sample SP-1074, and the cyan histogram corresponds to the $m_{A^0} = 2$ GeV Monte Carlo sample SP-8740. (a): number of crystals in the EmcBump; (b): LAT moment; (c): a_{42} moment; (d): $\cos\theta_\gamma^*$; (e): E_2; (f) E_{extra}.
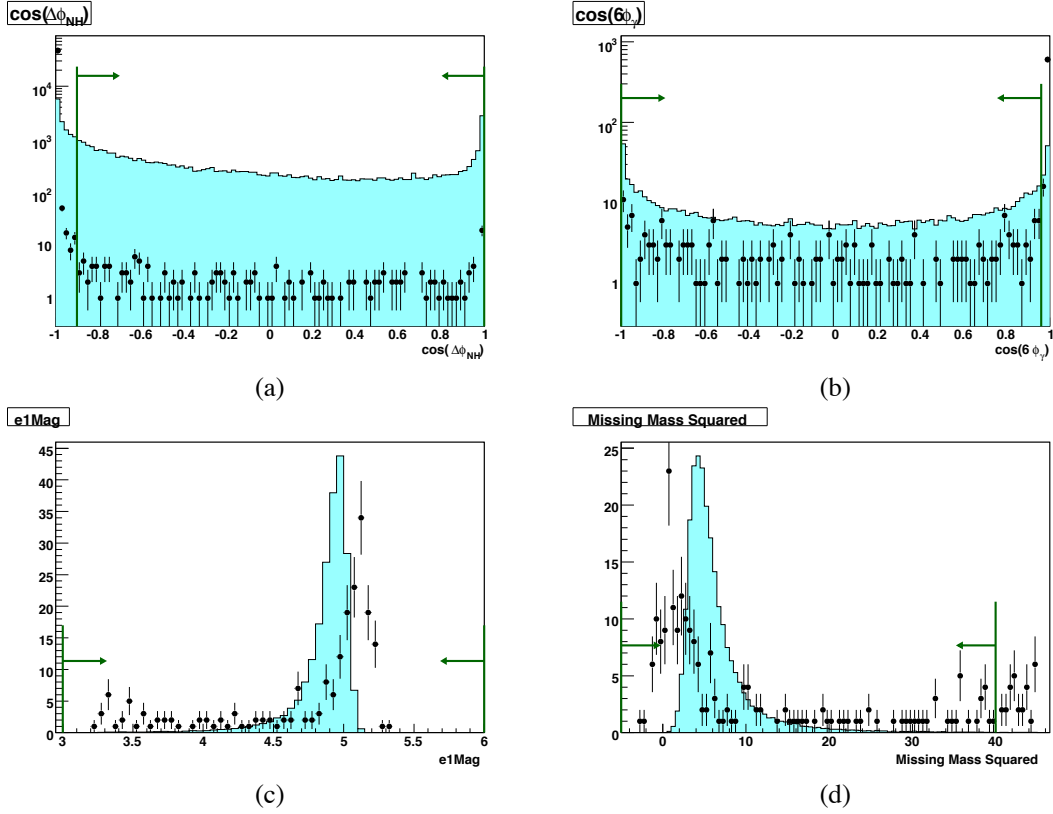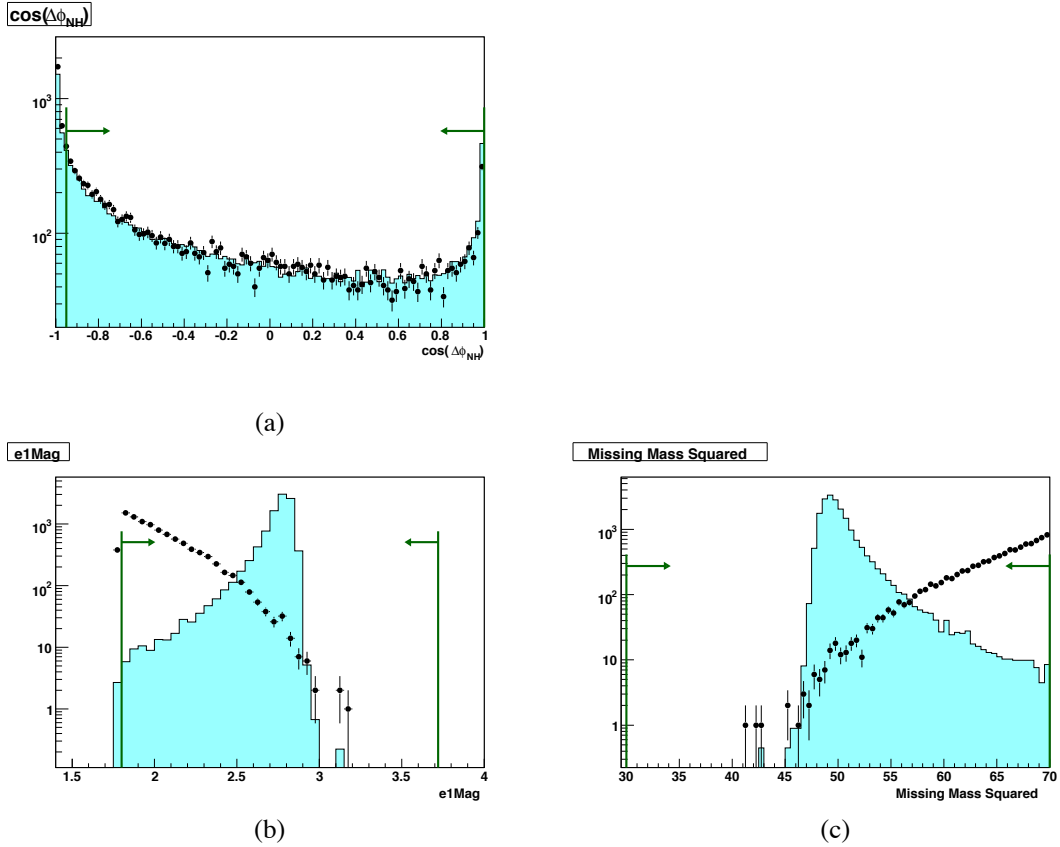
Figure 68: Plots of each of the selection variables with all final cuts, except on the variable plotted, applied. The selections for the high energy region, $3.2 < E^*_\gamma < 5.5$ are applied. The points represent the $e^+e^- \to \gamma\gamma$ sample SP-1074, and the cyan histogram corresponds to the $m_{A^0} = 2$ GeV Monte Carlo sample SP-8740. The green lines show the selection interval. (a): $\cos\Delta\phi_{NH}$; (b): $\cos(6\phi_\gamma)$; (c): $< E^*_\gamma$; (d): recoil mass squared m^2_X.
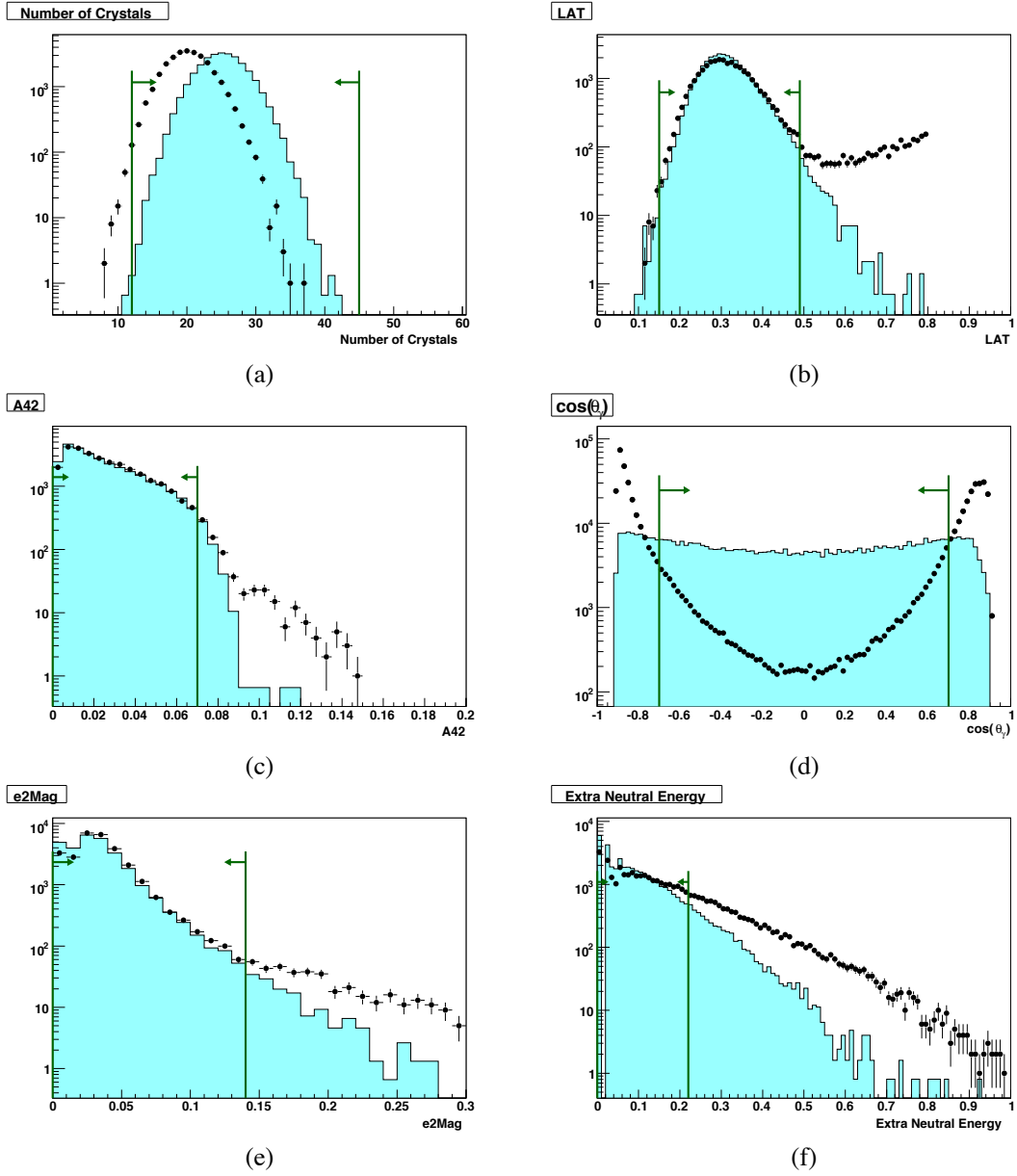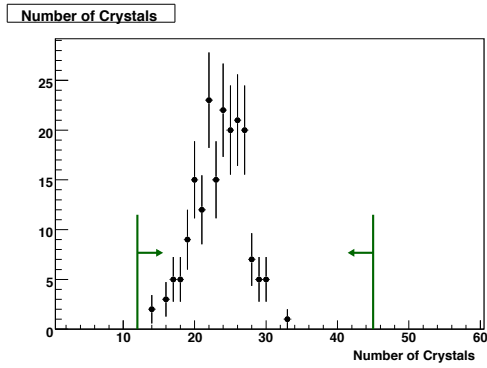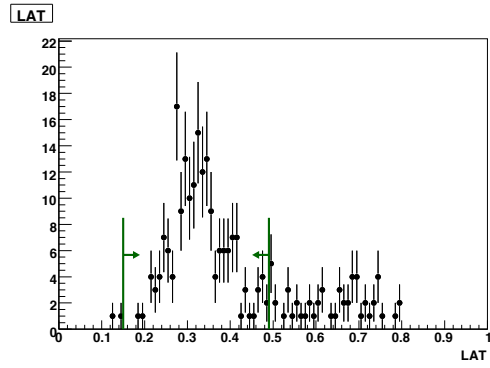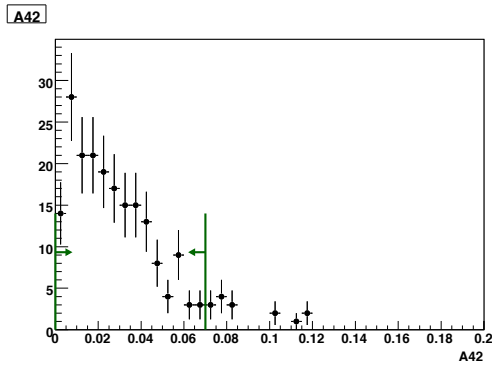
Figure 69: Plots of each of the selection variables with all final cuts, except on the variable plotted, applied. The selections for the low energy region, $1.8 < E_\gamma^* < 3.7$ are applied. The points represent the 2.4 fb^{-1} "BGFilterSkim-High" dataset, and the cyan histogram corresponds to the $m_{A^0} = 7$ GeV Monte Carlo sample SP-8921. (a): number of crystals in the EmcBump; (b): LAT moment; (c): a_{42} moment; (d): $\cos\theta_\gamma^*$; (e): E_2; (f) E_{extra}.

Figure 70: Plots of each of the selection variables with all final cuts, except on the variable plotted, applied. The selections for the low energy region, $1.8 < E_\gamma^* < 3.7$ are applied. The points represent the 2.4 fb^{-1} "BGFilterSkim-High" dataset, and the cyan histogram corresponds to the $m_{A^0} = 7$ GeV Monte Carlo sample SP-8921. The green lines show the selection interval. (a): $\cos\Delta\phi_{NH}$; (b): $< E_\gamma^*$; (c): recoil mass squared m_X^2.
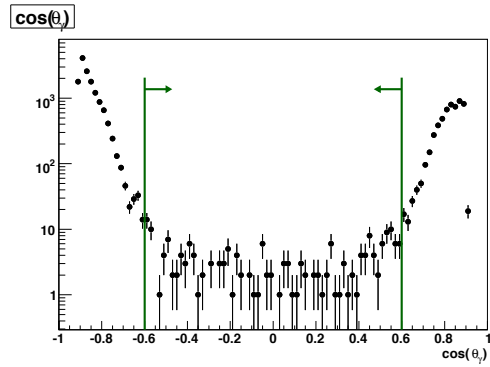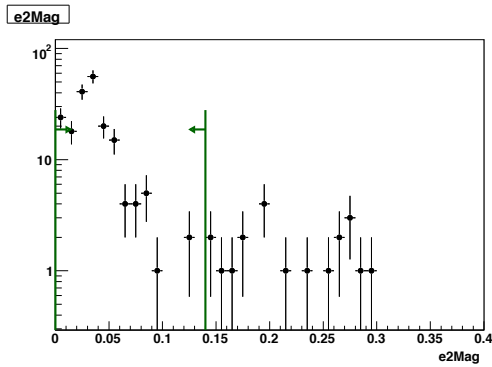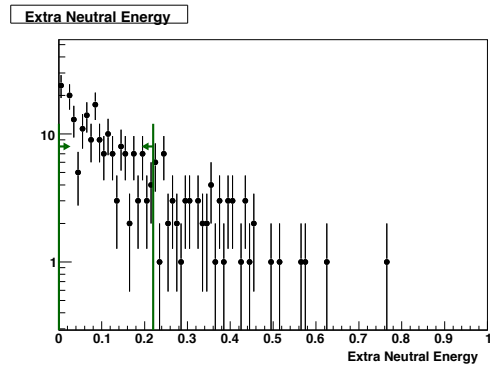
Figure 71: Plots of each of the selection variables with all final cuts, except on the variable plotted, applied. The selections for the low energy region, $1.8 < E_\gamma^* < 3.7$ are applied. The points represent the $e^+e^- \rightarrow \gamma\gamma$ sample `SP-1074`, and the cyan histogram corresponds to the $m_{A^0} = 7$ GeV Monte Carlo sample `SP-8921`. (a): number of crystals in the EmcBump; (b): LAT moment; (c): a_{42} moment; (d): $\cos\theta_\gamma^*$; (e): E_2; (f) E_{extra}.
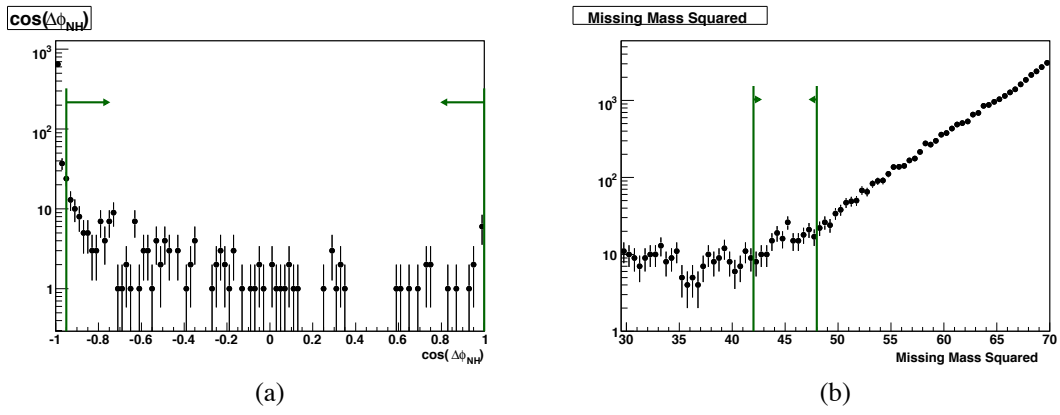
(a)

(b)   (c)

Figure 72: Plots of each of the selection variables with all final cuts, except on the variable plotted, applied. The selections for the low energy region, $1.8 < E_\gamma^* < 3.7$ are applied. The points represent the $e^+e^- \to \gamma\gamma$ sample SP-1074, and the cyan histogram corresponds to the $m_{A^0} = 7$ GeV Monte Carlo sample SP-8921. The green lines show the selection interval. (a): $\cos\Delta\phi_{NH}$; (b): $< E_\gamma^*$; (c): recoil mass squared m_X^2.

Figure 73: Plots of each of the selection variables with all final cuts, except on the variable plotted, applied. The selections for the low energy region, $1.8 < E_\gamma^* < 3.7$ are applied. The points represent the low-angle bhabha sample SP-8939, and the cyan histogram corresponds to the $m_{A^0} = 7$ GeV Monte Carlo sample SP-8921. (a): number of crystals in the EmcBump; (b): LAT moment; (c): a_{42} moment; (d): $\cos\theta_\gamma^*$; (e): E_2; (f) E_{extra}.

Figure 74: Plots of each of the selection variables with all final cuts, except on the variable plotted, applied. The selections for the low energy region, $1.8 < E_\gamma^* < 3.7$ are applied. The points represent the low-angle bhabha sample SP-8939, and the cyan histogram corresponds to the $m_{A^0} = 7$ GeV Monte Carlo sample SP-8921. The green lines show the selection interval. (a): $\cos \Delta\phi_{NH}$; (b): $< E_\gamma^*$; (c): recoil mass squared m_X^2.

Figure 75: Plots of each of the selection variables with all final cuts, except on the variable plotted, applied. The selections around the $m_{A^0} = 6.7$ GeV are applied: $42 < m_X^2 < 48$ GeV^2. The points represent the 2.4 fb^{-1} "BGFilterSkim-High" dataset. (a): number of crystals in the EmcBump; (b): LAT moment; (c): a_{42} moment; (d): $\cos\theta_\gamma^*$; (e): E_2; (f) E_{extra}.

Figure 76: Plots of each of the selection variables with all final cuts, except on the variable plotted, applied. The selections around the $m_{A^0} = 6.7$ GeV are applied: $42 < m_X^2 < 48$ GeV^2. The points represent the 2.4 fb^{-1} "BGFilterSkim-High" dataset. The green lines show the selection interval. (a): $\cos \Delta\phi_{NH}$; (b): recoil mass squared m_X^2.

B   Signal PDF Fits
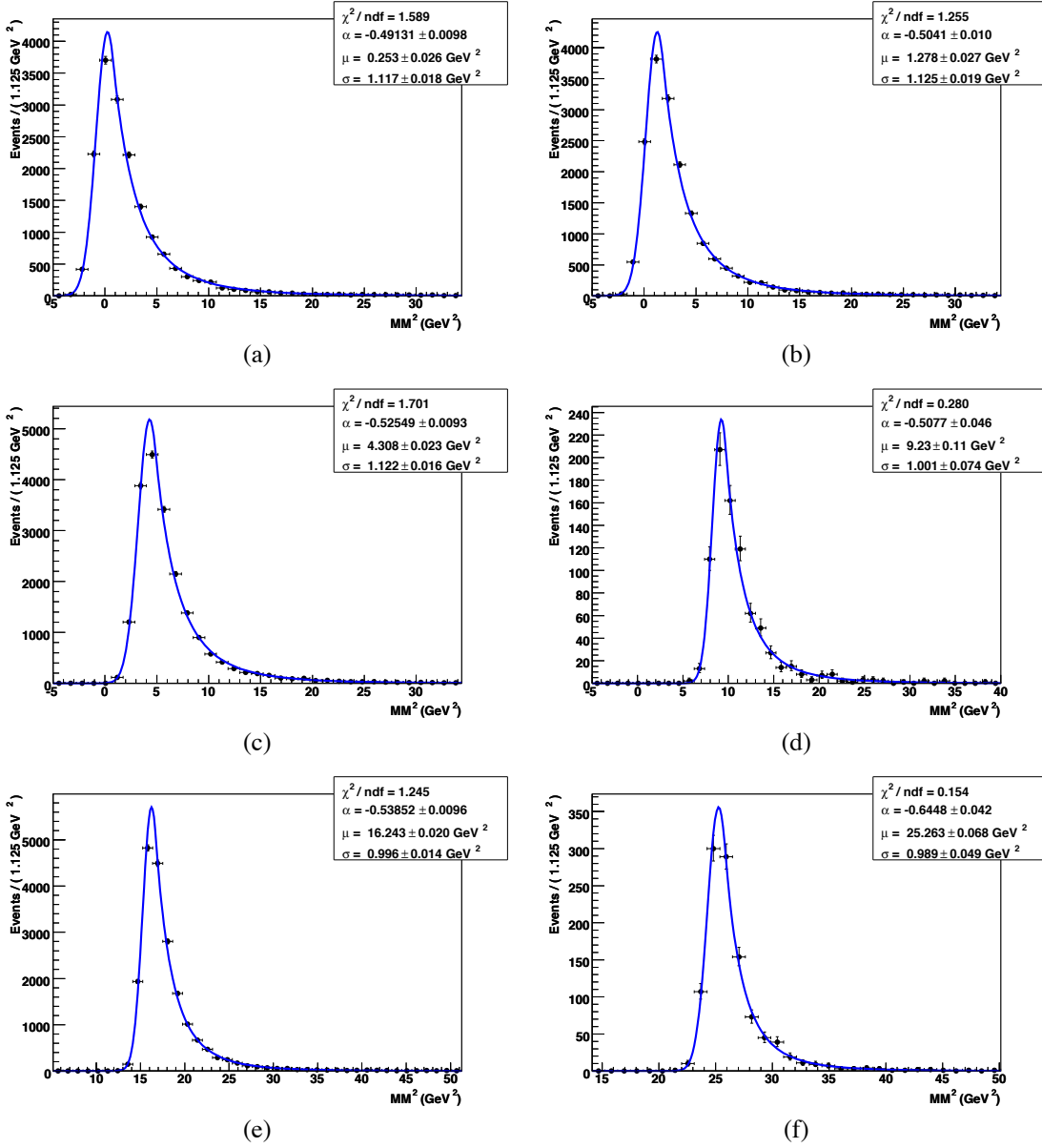
Figure 77: Fits to the m_X^2 distribution of signal $\Upsilon(3S) \rightarrow \gamma A^0$ events for assumed A^0 masses of (a) $m_{A^0} = 0.1$ GeV, (b) $m_{A^0} = 1.0$ GeV, (c) $m_{A^0} = 2.0$ GeV, (d) $m_{A^0} = 3.0$ GeV, (e) $m_{A^0} = 4.0$ GeV, (f) $m_{A^0} = 5.0$ GeV.

71
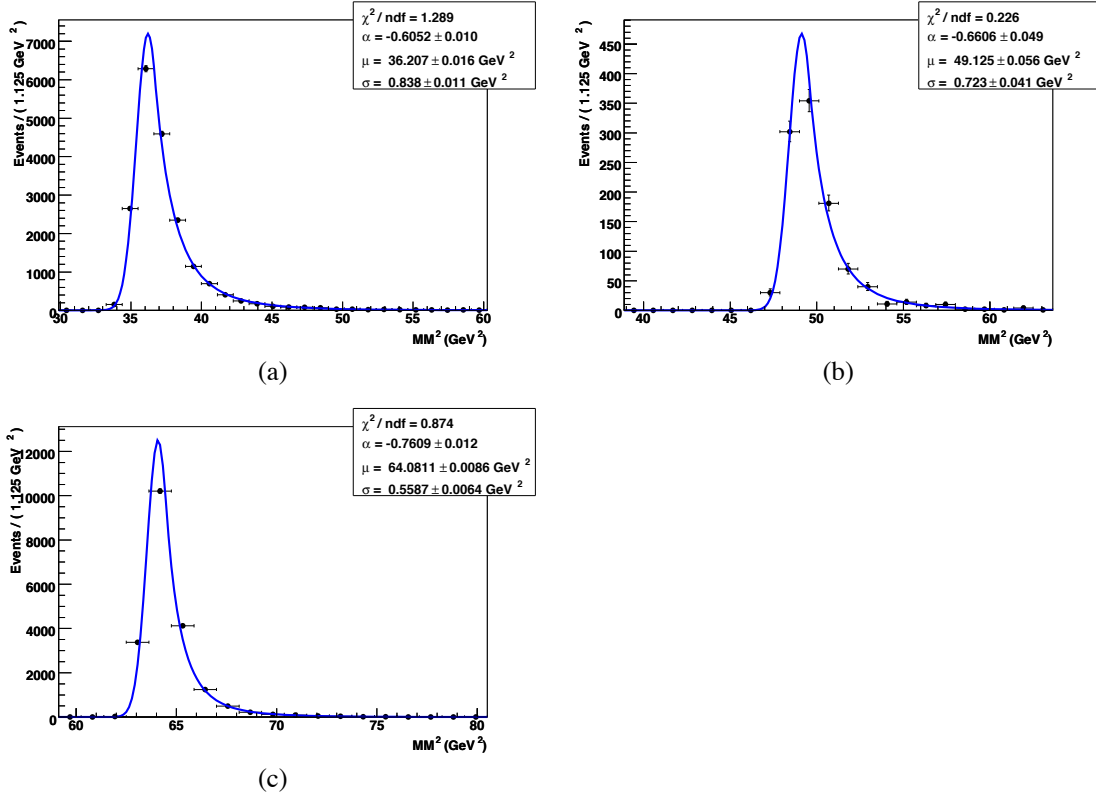
Figure 78: Fits to the m_X^2 distribution of signal $\Upsilon(3S) \rightarrow \gamma A^0$ events for assumed A^0 masses of (a) $m_{A^0} = 6.0$ GeV, (b) $m_{A^0} = 7.0$ GeV, (c) $m_{A^0} = 8.0$ GeV.

C   Results of the Toy Monte Carlo

We test the fitting procedure with the toy Monte Carlo fits. For each toy experiment, we generate the background from the PDF obtained in the data fit, and embed a known amount of signal (0 or 10) events from the Monte Carlo simulations. For the high-energy region, we also embed 10 $e^+e^- \rightarrow \gamma\gamma$ events from the data, selected without the IFR veto.

C.1   High energy region

We show results of the toy studies in the high energy region $m_{A^0} < 6$ GeV in Fig. 79-92. We fit each sample with a fixed value of assumed m_{A^0} from $m_{A^0} = 0.1$ GeV to $m_{A^0} = 6$ GeV. In general, with the exception of a few outliers from poorly converged fits, no ensemble of the pseudo-experiments show any significant bias.
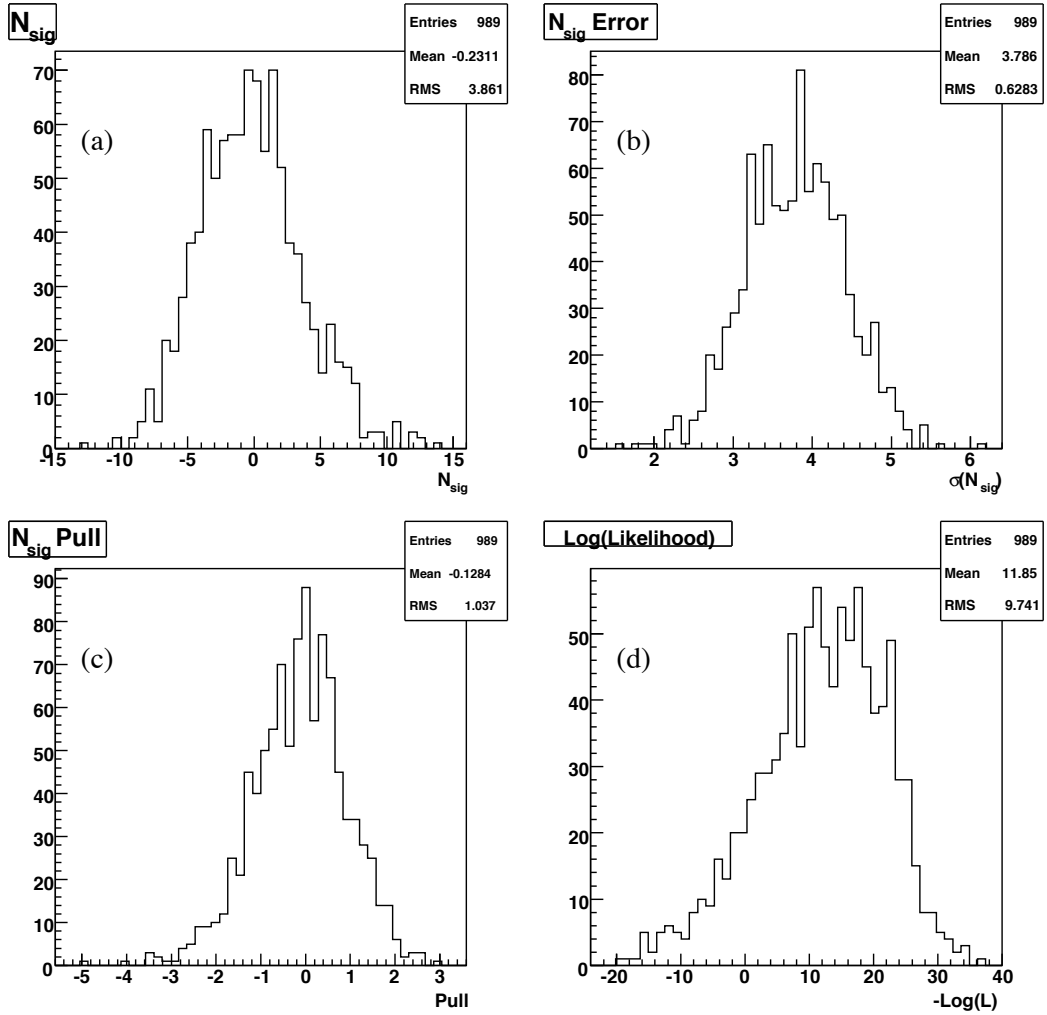
Figure 79: Toy Monte Carlo results for $N_{\mathrm{sig}} = 0$ and $m_{A^0} = 0.1$ GeV. (a): signal yield, (b): error, (c): pull, (d): likelihood value $-\ln(L)$.
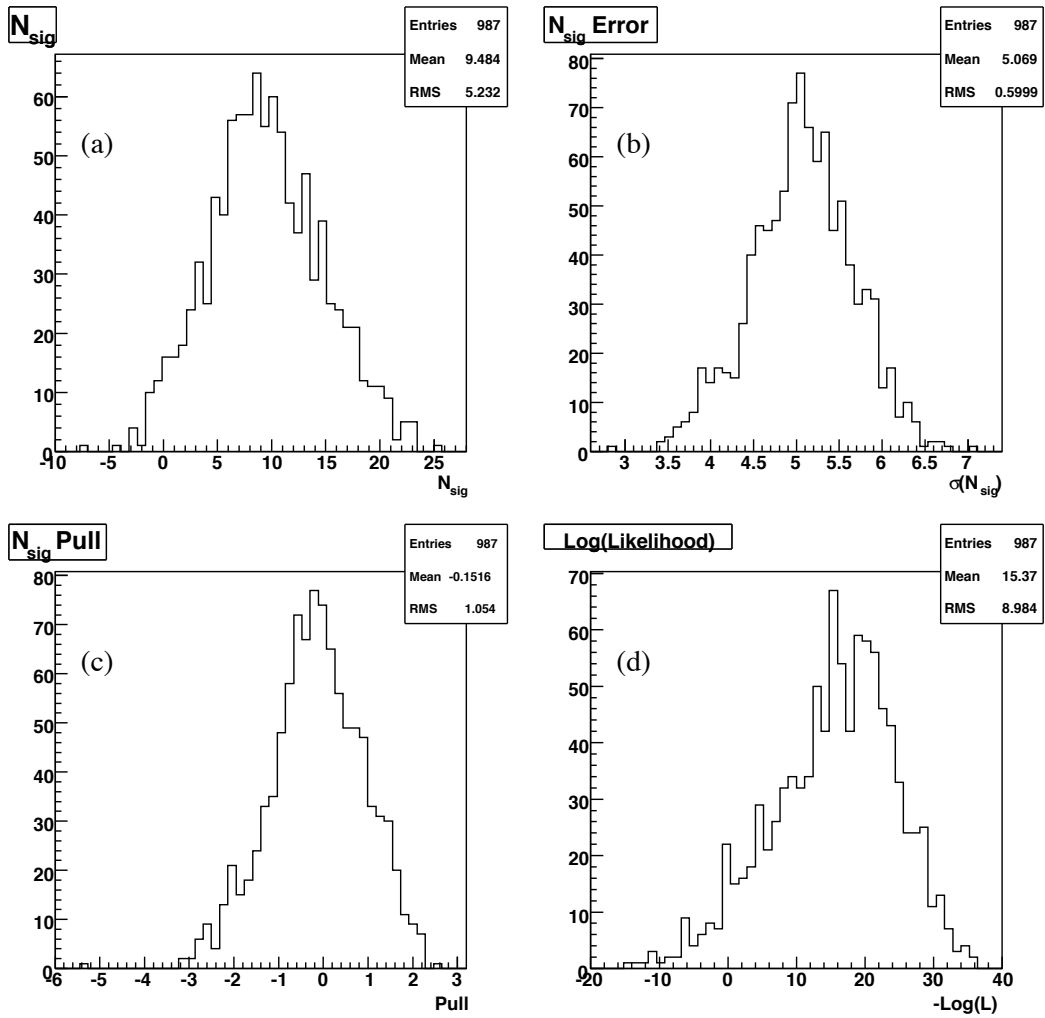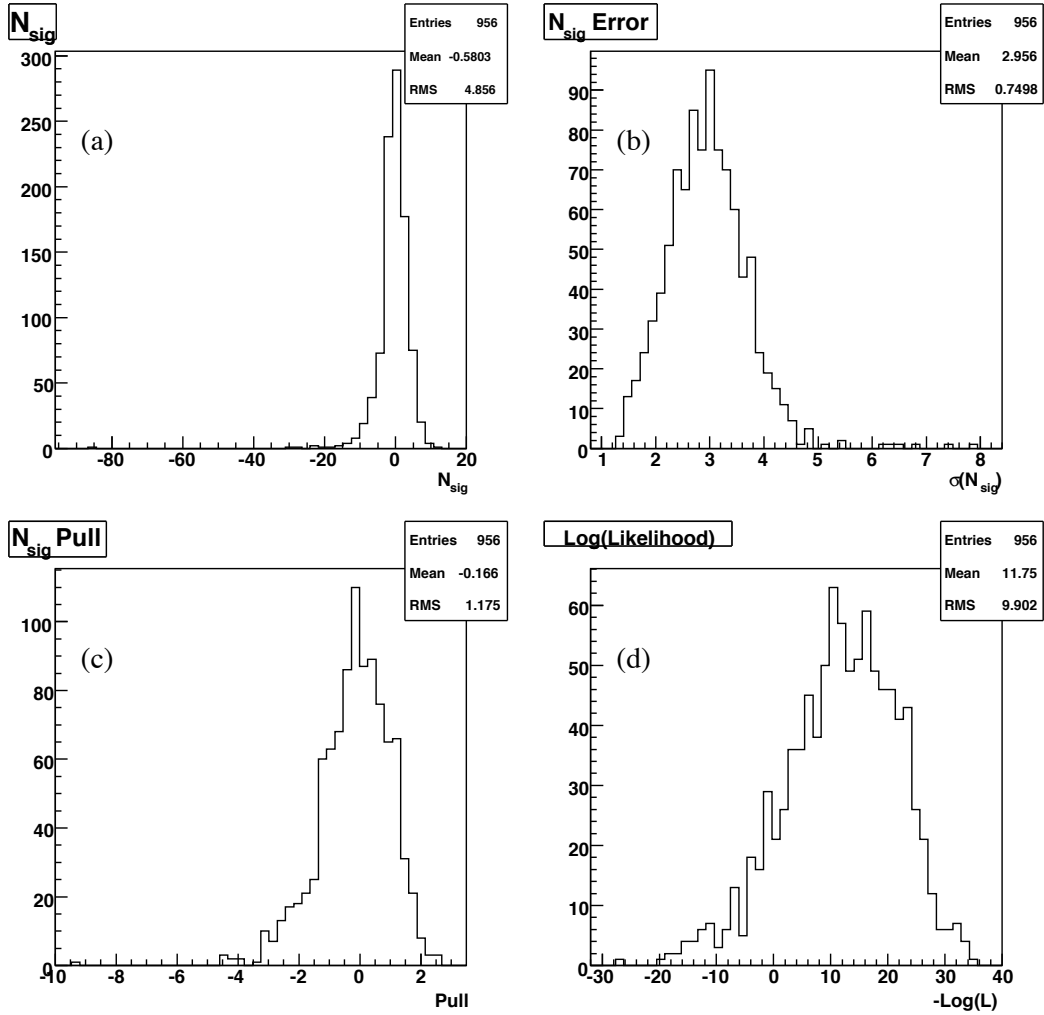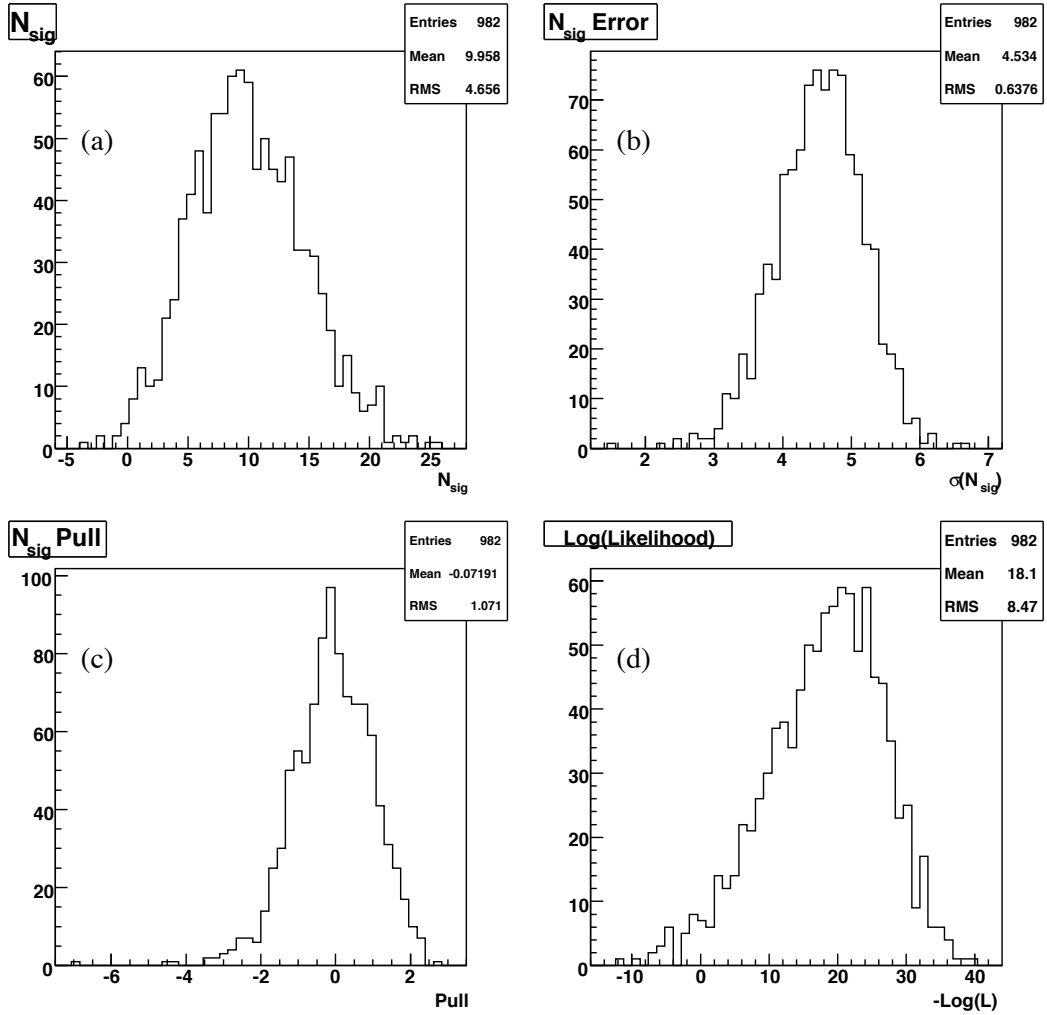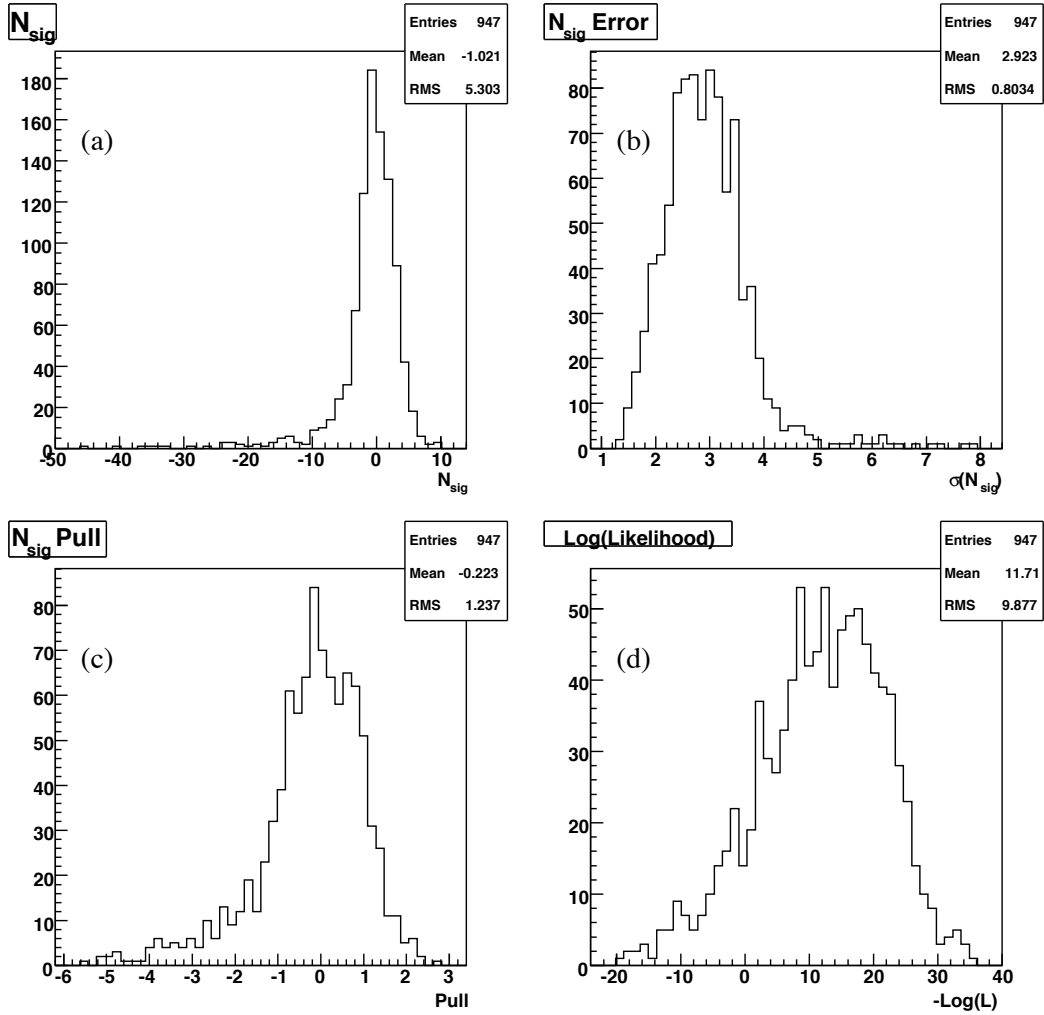
Figure 80: Toy Monte Carlo results for $N_{\mathrm{sig}} = 10$ and $m_{A^0} = 0.1$ GeV. (a): signal yield, (b): error, (c): pull, (d): likelihood value $-\ln(L)$.
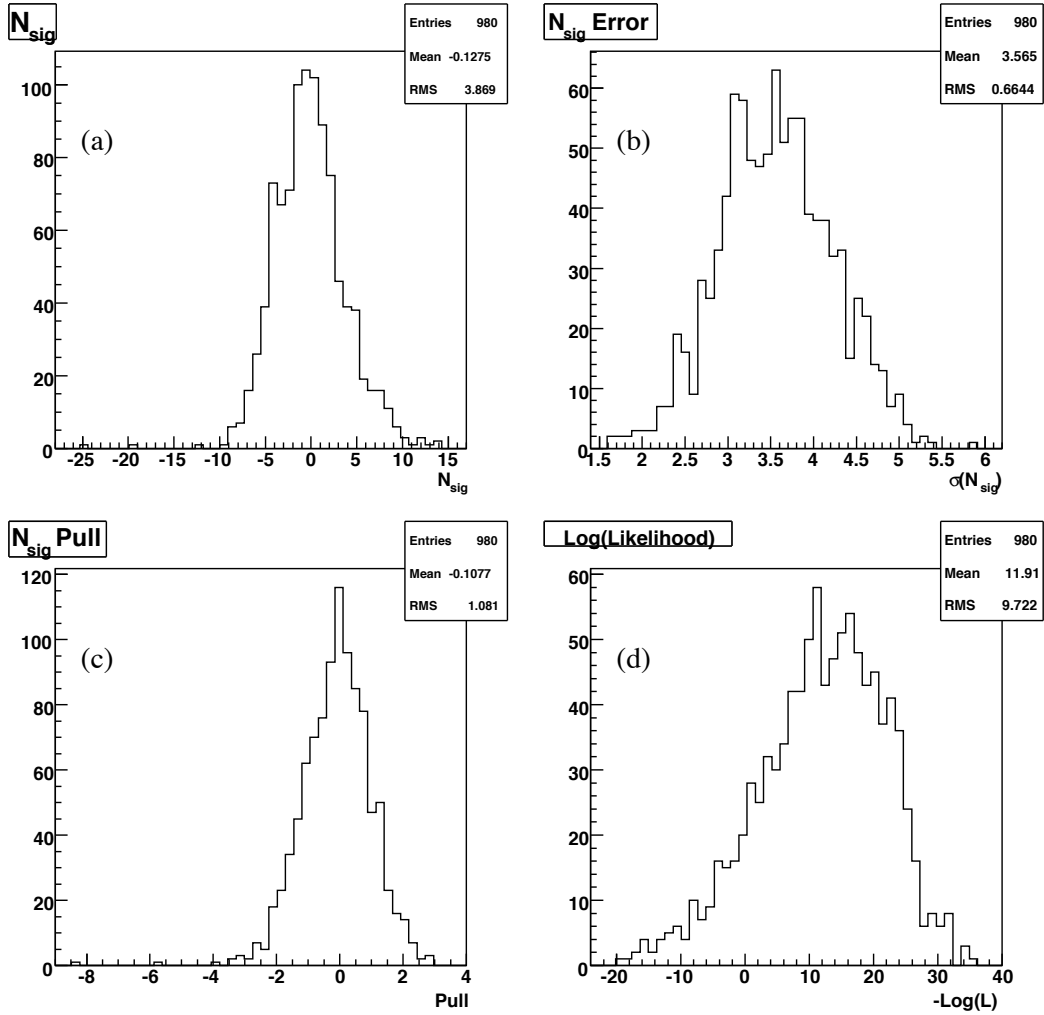
Figure 81: Toy Monte Carlo results for $N_{\text{sig}} = 0$ and $m_{A^0} = 1.0$ GeV. (a): signal yield, (b): error, (c): pull, (d): likelihood value $-\ln(L)$.
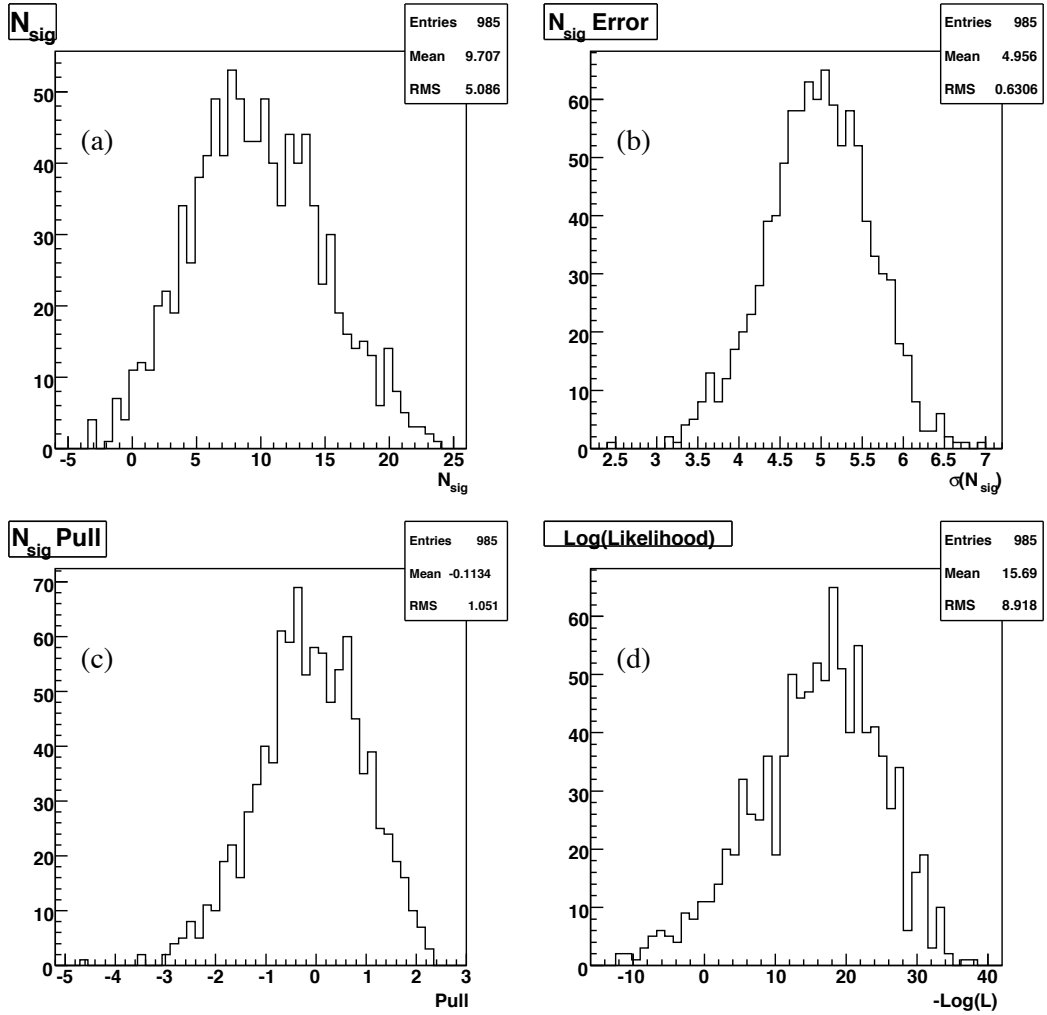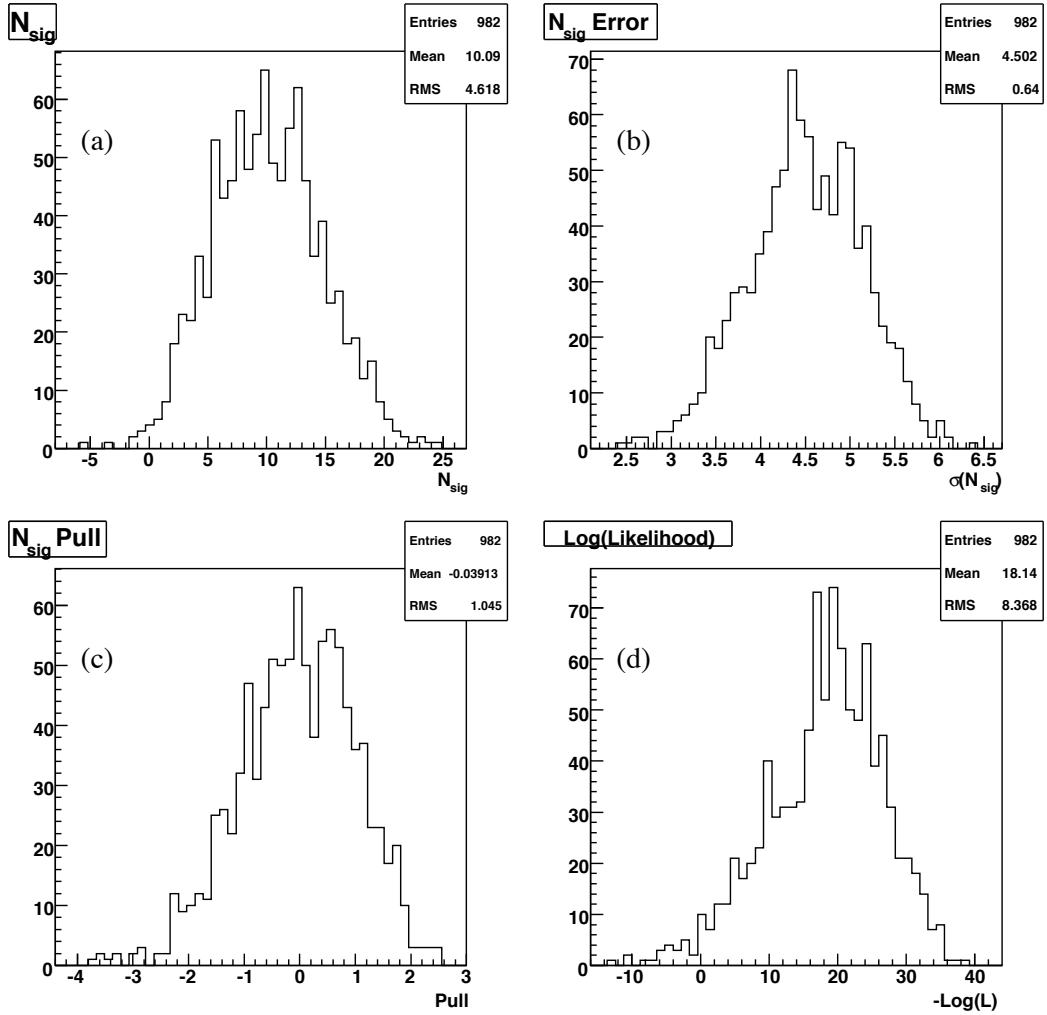
Figure 82: Toy Monte Carlo results for $N_{\text{sig}} = 10$ and $m_{A^0} = 1.0$ GeV. (a): signal yield, (b): error, (c): pull, (d): likelihood value $-\ln(L)$.
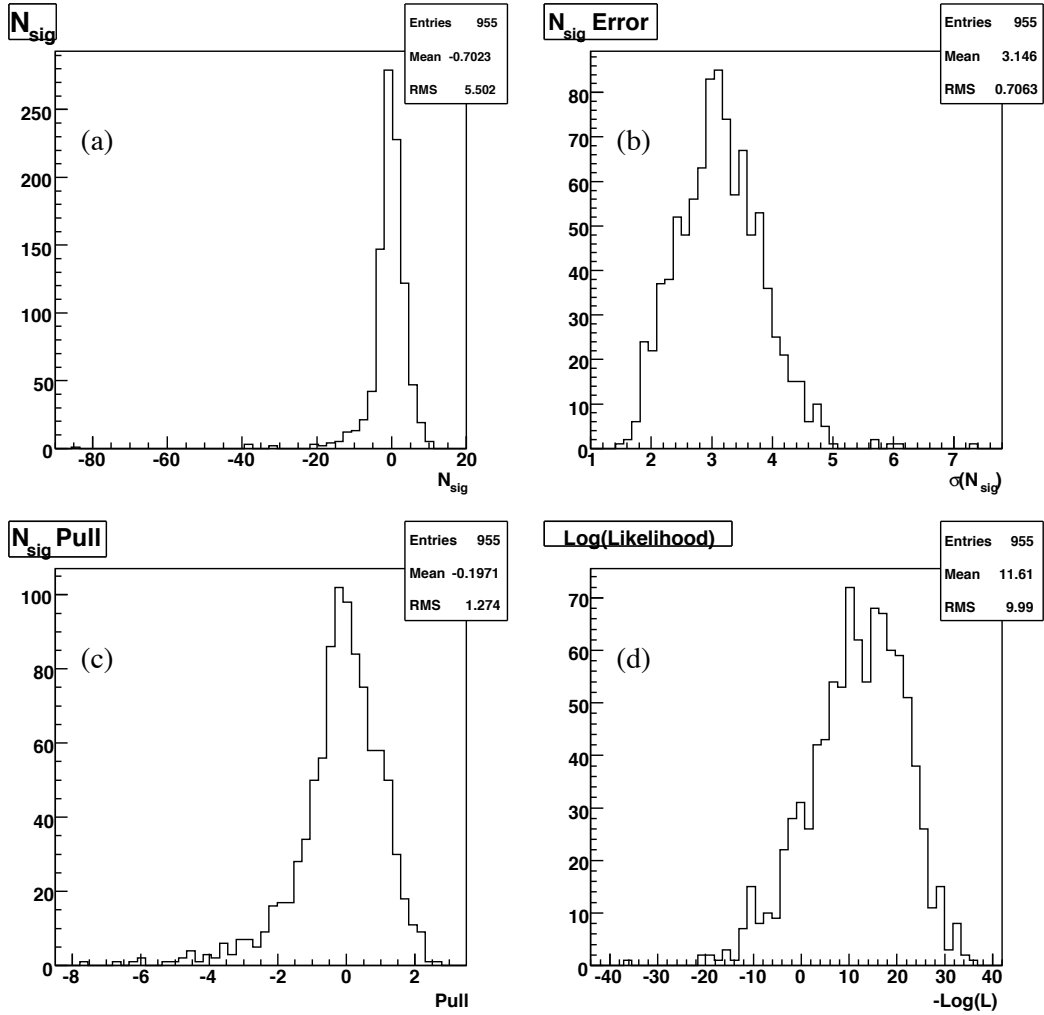
Figure 83: Toy Monte Carlo results for $N_{\mathrm{sig}} = 0$ and $m_{A^0} = 2.0$ GeV. (a): signal yield, (b): error, (c): pull, (d): likelihood value $-\ln(L)$.

Figure 84: Toy Monte Carlo results for $N_{\text{sig}} = 10$ and $m_{A^0} = 2.0$ GeV. (a): signal yield, (b): error, (c): pull, (d): likelihood value $-\ln(L)$.

Figure 85: Toy Monte Carlo results for $N_{\text{sig}} = 0$ and $m_{A^0} = 3.0$ GeV. (a): signal yield, (b): error, (c): pull, (d): likelihood value $-\ln(L)$.
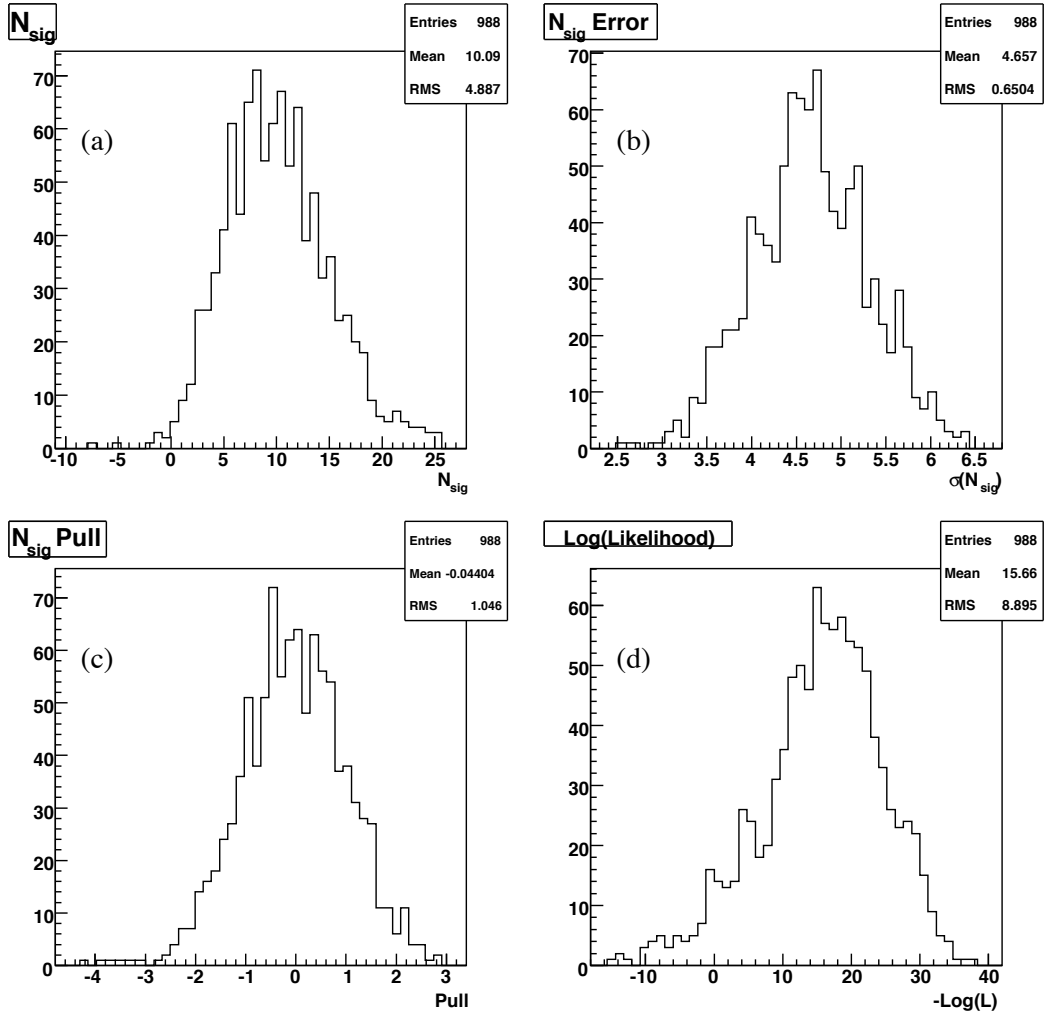
Figure 86: Toy Monte Carlo results for $N_{\mathrm{sig}} = 10$ and $m_{A^0} = 3.0$ GeV. (a): signal yield, (b): error, (c): pull, (d): likelihood value $-\ln(L)$.

Figure 87: Toy Monte Carlo results for $N_{\mathrm{sig}} = 0$ and $m_{A^0} = 4.0$ GeV. (a): signal yield, (b): error, (c): pull, (d): likelihood value $-\ln(L)$.
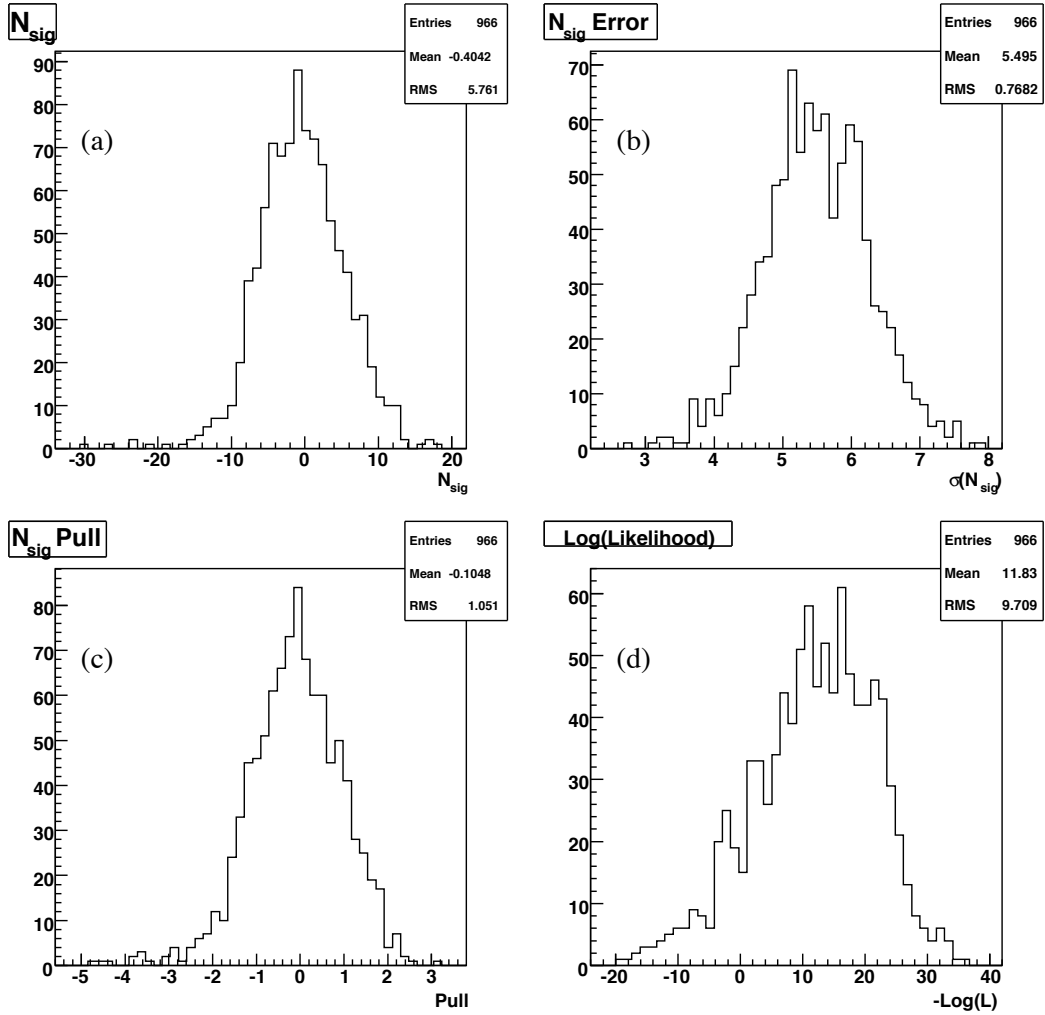
Figure 88: Toy Monte Carlo results for $N_{sig} = 10$ and $m_{A^0} = 4.0$ GeV. (a): signal yield, (b): error, (c): pull, (d): likelihood value $-\ln(L)$.
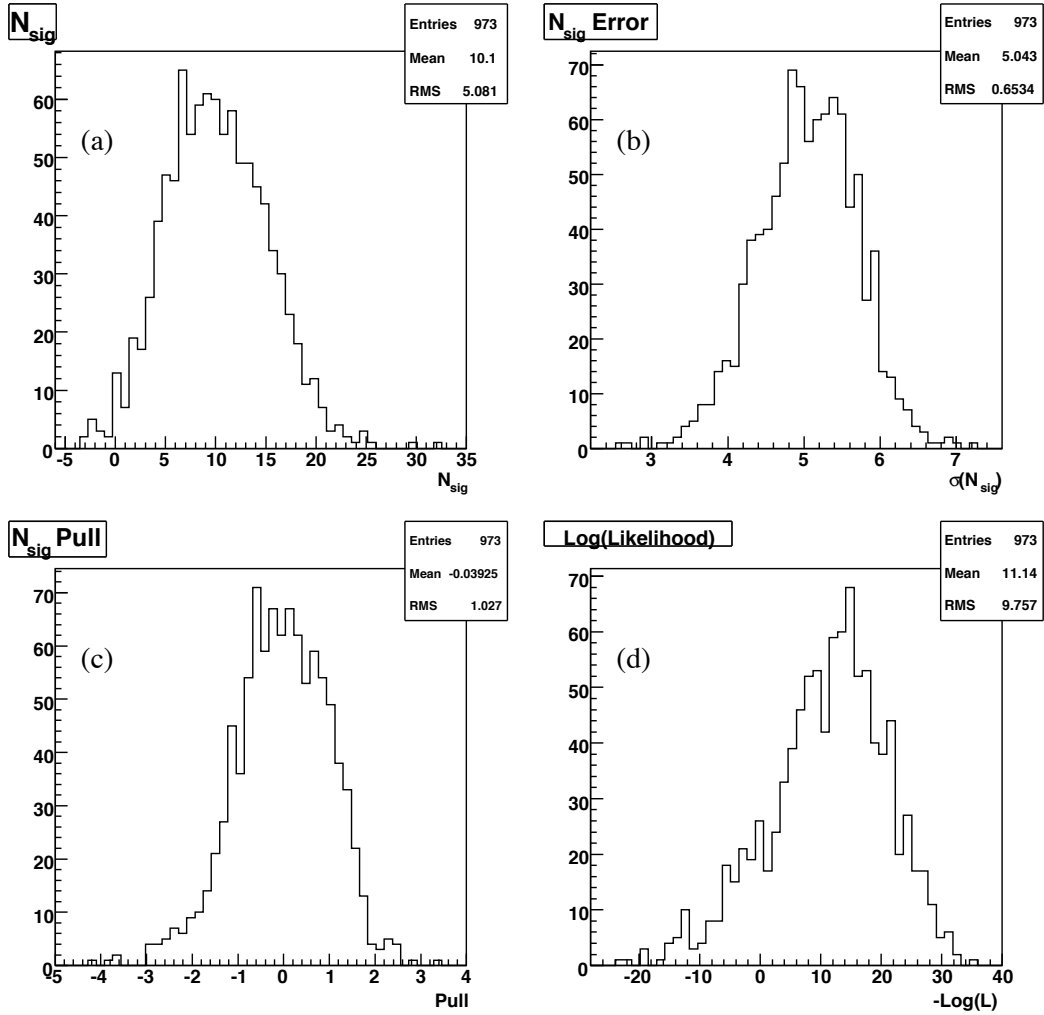
Figure 89: Toy Monte Carlo results for $N_{\text{sig}} = 0$ and $m_{A^0} = 5.0$ GeV. (a): signal yield, (b): error, (c): pull, (d): likelihood value $-\ln(L)$.

Figure 90: Toy Monte Carlo results for $N_{\text{sig}} = 10$ and $m_{A^0} = 5.0$ GeV. (a): signal yield, (b): error, (c): pull, (d): likelihood value $-\ln(L)$.
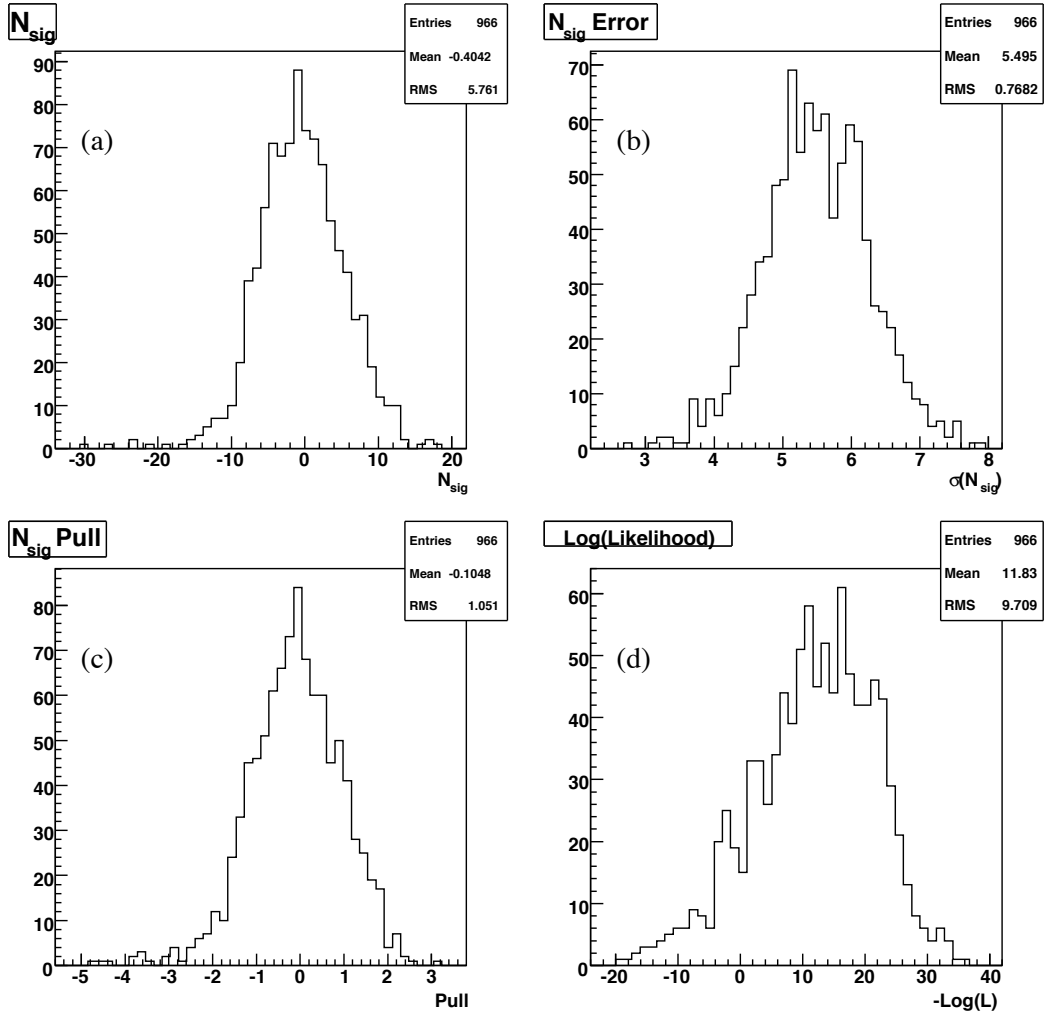
Figure 91: Toy Monte Carlo results for $N_{\mathrm{sig}} = 0$ and $m_{A^0} = 6.0$ GeV. (a): signal yield, (b): error, (c): pull, (d): likelihood value $-\ln(L)$.
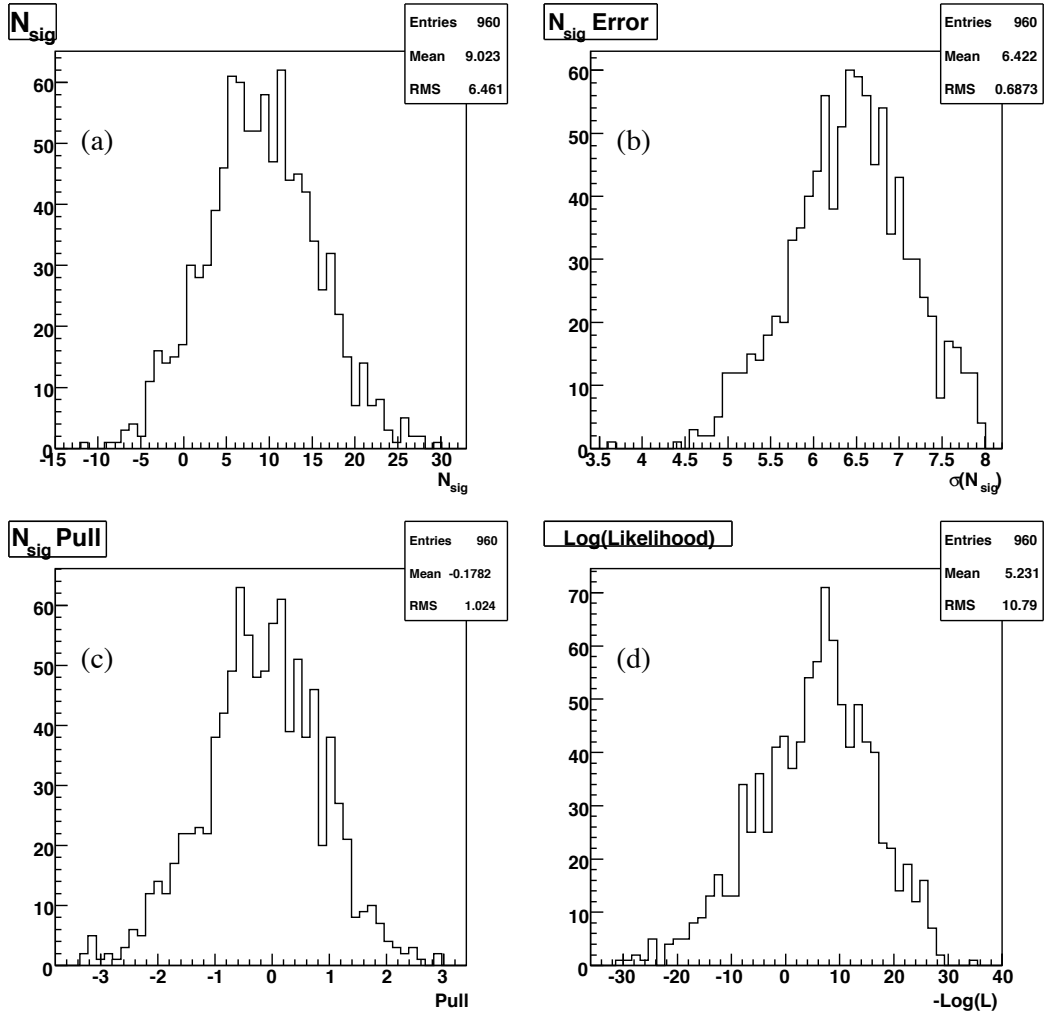
Figure 92: Toy Monte Carlo results for $N_{\mathrm{sig}} = 10$ and $m_{A^0} = 6.0$ GeV. (a): signal yield, (b): error, (c): pull, (d): likelihood value $-\ln(L)$.