First Look at 2 $\mathbf{ab^{-1}}$ SM MC Data Sample at $\sqrt{s} = 1$ TeV

August 13, 2002

Tim Barklow

Monte Carlo Production

WHIZARD is used to generate all of the SM processes $e^+e^- \rightarrow f_1f_2$, $f_1f_2f_3f_4$, and $f_1f_2f_3f_4f_5f_6$ including ISR & beamstrahlung (CIRCE).

Goal is to generate 2000 fb^{-1} MC data at $\sqrt{s} = 0.5$, 0.8, 1.0, 1.2, and 1.5 TeV.

100% electron *and* positron polarization is always assumed in event generation. Arbitrary electron/positron polarization is simulated by combining e^-_L/e^+_R, e^-_R/e^+_L, ... data sets.

Fully fragmented MC data sets are produced. PYTHIA is used for final state QED and QCD parton showering, as well as for fragmention and decay.

SM Final States 0,2,4-Fermion

0-fermion

$e^+e^- \longrightarrow$	$\gamma\gamma$	
	$\gamma\gamma\gamma$	
	$\gamma\gamma\gamma\gamma$	
	$\gamma\gamma\gamma\gamma\gamma$	

2-fermion

$e^+e^- \longrightarrow$	ff	$f \neq \nu$
	$\nu\nu\gamma$	
	$\nu\nu\gamma\gamma$	
	$\nu\nu\gamma\gamma\gamma$	
$e^-\gamma \longrightarrow$	$e^-\gamma$	
$\gamma e^+ \longrightarrow$	$e^+\gamma$	

4-fermion

$e^+e^- \longrightarrow$	$\nu\nu\nu\nu\gamma$	6 total
	$u_j\overline{d}_j d_k\overline{u}_k$	25 total
		$\nu_e e^+ e^- \overline{\nu}_e$
		$\nu_e e^+ \mu^- \overline{\nu}_\mu$
		$\nu_e e^+ \tau^- \overline{\nu}_\tau$
		$\nu_e e^+ d\overline{u}$
		.
		.
		$c\overline{s}s\overline{c}$
	$u_j\overline{u}_j u_k\overline{u}_k$	9 total
	$u_j\overline{u}_j d_k\overline{d}_k$	25 total
	$d_j\overline{d}_j d_k\overline{d}_k$	21 total
$\gamma\gamma \longrightarrow$	$f\overline{f}$	8 total
$e_L^-\gamma \longrightarrow$	$\nu_e d_k\overline{u}_k$	5 total
$e^-\gamma \longrightarrow$	$e^- f\overline{f}$	10 total
$\gamma e_R^+ \longrightarrow$	$\overline{\nu}_e u_k\overline{d}_k$	5 total
$\gamma e^+ \longrightarrow$	$e^+ f\overline{f}$	10 total

SM Final States 6-Fermion

6-fermion

$e^+e^- \longrightarrow$	$u_i\overline{u}_iu_j\overline{d}_jd_k\overline{u}_k$	125 total
	$d_i\overline{d}_iu_j\overline{d}_jd_k\overline{u}_k$	150 total
	$u_i\overline{u}_iu_j\overline{u}_ju_k\overline{u}_k$	25 total
	$u_i\overline{u}_iu_j\overline{u}_jd_k\overline{d}_k$	65 total
	$u_i\overline{u}_id_j\overline{d}_jd_k\overline{d}_k$	75 total
	$d_i\overline{d}_id_j\overline{d}_jd_k\overline{d}_k$	56 total
$\gamma\gamma \longrightarrow$	$u_j\overline{d}_jd_k\overline{u}_k$	25 total
	$u_j\overline{u}_ju_k\overline{u}_k$	9 total
	$u_j\overline{u}_jd_k\overline{d}_k$	25 total
	$d_j\overline{d}_jd_k\overline{d}_k$	21 total
$e_L^-\gamma \longrightarrow$	$\nu_eu_j\overline{u}_jd_k\overline{u}_k$	25 total
	$\nu_ed_j\overline{d}_jd_k\overline{u}_k$	30 total
$e^-\gamma \longrightarrow$	$e^-u_j\overline{d}_jd_k\overline{u}_k$	20 total
	$e^-u_j\overline{u}_ju_k\overline{u}_k$	10 total
	$e^-u_j\overline{u}_jd_k\overline{d}_k$	20 total
	$e^-d_j\overline{d}_jd_k\overline{d}_k$	21 total
$\gamma e_R^+ \longrightarrow$	$\overline{\nu}_eu_j\overline{d}_ju_k\overline{u}_k$	25 total
	$\overline{\nu}_eu_j\overline{d}_jd_k\overline{d}_k$	30 total
$\gamma e^+ \longrightarrow$	$e^+u_j\overline{d}_jd_k\overline{u}_k$	20 total
	$e^+u_j\overline{u}_ju_k\overline{u}_k$	10 total
	$e^+u_j\overline{u}_jd_k\overline{d}_k$	20 total
	$e^+d_j\overline{d}_jd_k\overline{d}_k$	21 total

SM Final States 8-Fermion

8-fermion

$e^+e^- \longrightarrow$	$f\overline{f}t\overline{t}$
$\gamma\gamma \longrightarrow$	$t\overline{t}$
$e^-\gamma \longrightarrow$	$e^-t\overline{t}$
	$\nu_e b\overline{t}$
$\gamma e^+ \longrightarrow$	$e^+t\overline{t}$
	$\overline{\nu}_e t\overline{b}$

WHIZARD MC uses the CompHEP convention for particle names, and we use them as well when specifying processes:

```
St.Model(Feyn.gauge)
 Particles
 Full  name   | P | aP|2*spin| mass |width |color|aux|
photon       |A  |A  |2     |0     |0     |1    |G
gluon        |G  |G  |2     |0     |0     |8    |G
electron     |e1 |E1 |1     |0     |0     |1    |
e-neutrino   |n1 |N1 |1     |0     |0     |1    |L
muon         |e2 |E2 |1     |Mm    |0     |1    |
m-neutrino   |n2 |N2 |1     |0     |0     |1    |L
tau-lepton   |e3 |E3 |1     |Mt    |0     |1    |
t-neutrino   |n3 |N3 |1     |0     |0     |1    |L
u-quark      |u  |U  |1     |0     |0     |3    |
d-quark      |d  |D  |1     |0     |0     |3    |
c-quark      |c  |C  |1     |Mc    |0     |3    |
s-quark      |s  |S  |1     |Ms    |0     |3    |
t-quark      |t  |T  |1     |Mtop  |wtop  |3    |
b-quark      |b  |B  |1     |Mb    |0     |3    |
Higgs        |H  |H  |0     |MH    |wH    |1    |
W-boson      |W+ |W- |2     |MW    |wW    |1    |G
Z-boson      |Z  |Z  |2     |MZ    |wZ    |1    |G
```

As a first test of interfacing a detector simulation program to this MC data set I used the ECFA/DESY fast MC program SIMDET. I made small modifications to only 1 subroutine in SIMDET:

```
timb@born2 $ diff sipyth.F $A6F/Simdet/simdet/code_f/sipyth.F
56,63d55
<
<       if(ifpyth.eq.3) then
<
<       call upinit(nevent,decm)
<       ecms=decm
<
<       else
<
321,323d312
<
<       endif
<
334,342d322
<
<          if(ifpyth.eq.3) then
<             call upevnt
<             IF (IFBKGR.NE.0 .AND. BKGEVT.GT.0.)  THEN    ! plus background
<                CALL HADES
<             END IF
<
<           else
<
421,422d400
<          end if
<
476,477d453
<       if(ifpyth.eq.3) call upprt
<
```

Polarization, $\sqrt{s}$, and specific processes are defined in whizdata.in:

```
 &whizdata_input
 path_root = '/afs/slac.stanford.edu/g/nld/whizard/'
 data_root = '/nfs/mstore/g/lcddata/'
 i_sqrts = 1000
 luminosity = 2000.
 n_events_max=120000
 mbyte_max = 200.
 pol_eminus = -1.0
 pol_eplus  =  1.0
 seed = 520027
 output_events = F
 process =
   "e1,E1    q,q,q,q"
   "e1,E1    l,q,l,q"
   "e1,E1    l,v,l,v,q,q"
   "e1,A     f,l,l,q,q"
   "e1,A     e1,e1,E1,e2,E2"
 /
```

where q,l,v,f,x are defined as:

```
q=u,d,s,c,b,U,D,S,C,B
l=e1,e2,e3,E1,E2,E3
v=n1,n2,n3,N1,N2,N3
f=q,l,v
x=f,A
e3=e3,E3
E3=E3
```

To read out all generated MC data:

```
 &whizdata_input
 path_root = '/afs/slac.stanford.edu/g/nld/whizard/'
 data_root = '/nfs/mstore/g/lcddata/'
 i_sqrts = 1000
 luminosity = 3000.
 pol_eminus = -0.8
 pol_eplus  =  0.5
 seed = 520033
 process =
   "x,x    x,x"
   "x,x    x,x,x"
   "x,x    x,x,x,x"
   "x,x    x,x,x,x,x"
   "x,x    x,x,x,x,x,x"
 /
```

To read out $t\bar{t}$ events:

```
&whizdata_input
path_root = '/afs/slac.stanford.edu/g/nld/whizard/'
data_root = '/nfs/mstore/g/lcddata/'
i_sqrts = 1000
luminosity = 3000.
pol_eminus = -0.8
pol_eplus  =  0.5
seed = 520033
process =
  "e1,E1    b,b,f,f,f,f"
/
```

To read out ZH events:

```
&whizdata_input
path_root = '/afs/slac.stanford.edu/g/nld/whizard/'
data_root = '/nfs/mstore/g/lcddata/'
i_sqrts = 1000
luminosity = 3000.
pol_eminus = -0.8
pol_eplus  =  0.5
seed = 520033
process =
  "e1,E1    b,b,f,f"
  "e1,E1    e3,e3,f,f"
/
```

MC Limitations

- **Improper treatment of final states with identical quarks such as $u\bar{u}d\bar{d}$.** This is a parton-level generation problem as well as a fragmentation problem, and is due to a deficiency in the OMEGA amplitude calculation program. Eventually it will be fixed by the OMEGA author; in the meantime the parton-level problem can be corrected in principle with reweighting. The size of the effect is unknown, but should only be an issue in non-resonant phase space regions.
- **Events are not completely unweighted.** There is a wide range in cross sections, and we cannot, for example, generate an equivalent number of Bhabha's if we generate 2000 fb^{-1} of $e^+e^- \rightarrow u\bar{u}d\bar{d}s\bar{s}$. If you restrict yourself to high-pt central events, then the event sample will be unweighted. However, as you go out to the forward region you will eventually encounter some events with weight greater than 1 (bhabhas, $\gamma\gamma$ events). Thus one should always consider the event weight which is returned in PYTHIA variable PARI(7) in common PYPARS.
- **Top quark final state spin correlations missing in 8-fermion production.** WHIZARD is used to generate processes such as $e^+e^- \rightarrow f\bar{f}t\bar{t}$ without t quark decay. Hence final state spin correlations are missing. Also the t quark has 0 width.

SM MC Production Status as of June 18 2002

- Except for some very high cross section process, event generation for 0-2-4 fermion processes at $\sqrt{s} =$ 1000 GeV is complete.
- Integration of 6 fermion processes at $\sqrt{s} = 1000$ GeV is nearly complete.
- MC data sets are currently stored on MSTORE mass storage, with estimated 1.5 Terabytes storage per $\sqrt{s}$ point. External pythia process will handle all the MSTORE details for you.
- Good progress on writing code for external pythia process that interprets input file and reads out MC data. Should be ready about a week after Santa Cruz workshop.

SM MC Production Status as of August 13 2002

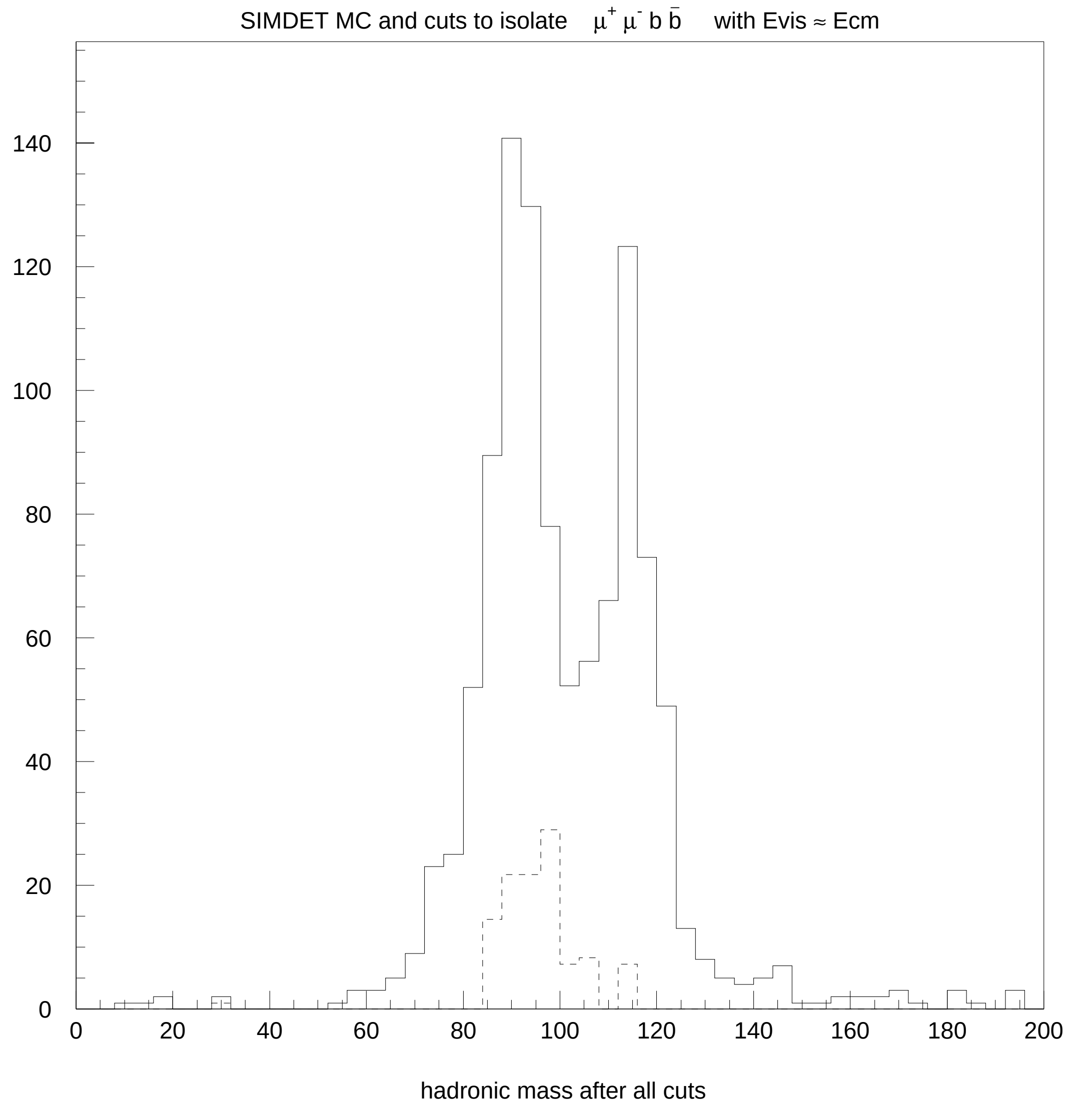
- Event generation for all 0-2-4-6 fermion processes at $\sqrt{s} = 1000$ GeV is complete.
- 8-fermion production is being held up by some complications involving t-quark decay in PYTHIA, but these issues will be resolved shortly.
- MC data sets are currently stored on MSTORE mass storage, arranged in 7216 stdhep files with a grand total storage of 3.05 Terabytes for the $\sqrt{s} = 1$ TeV data set. 0-2-4-fermion processes are contained in 2207 files and use 2.85 Terabytes while 6-fermion processes are contained in 5009 files and use 0.20 Terabytes.
- Code to interpret whizdata.in input file and to read out MC data seems to be working and the system has been successfully interfaced to the ECFA/DESY fast MC program SIMDET. It should be relatively straightforward to interface the data set to other detector simulation programs and analysis packages.

Sample analysis of $e^+e^- \to \mu^+\mu^- b\bar{b}$

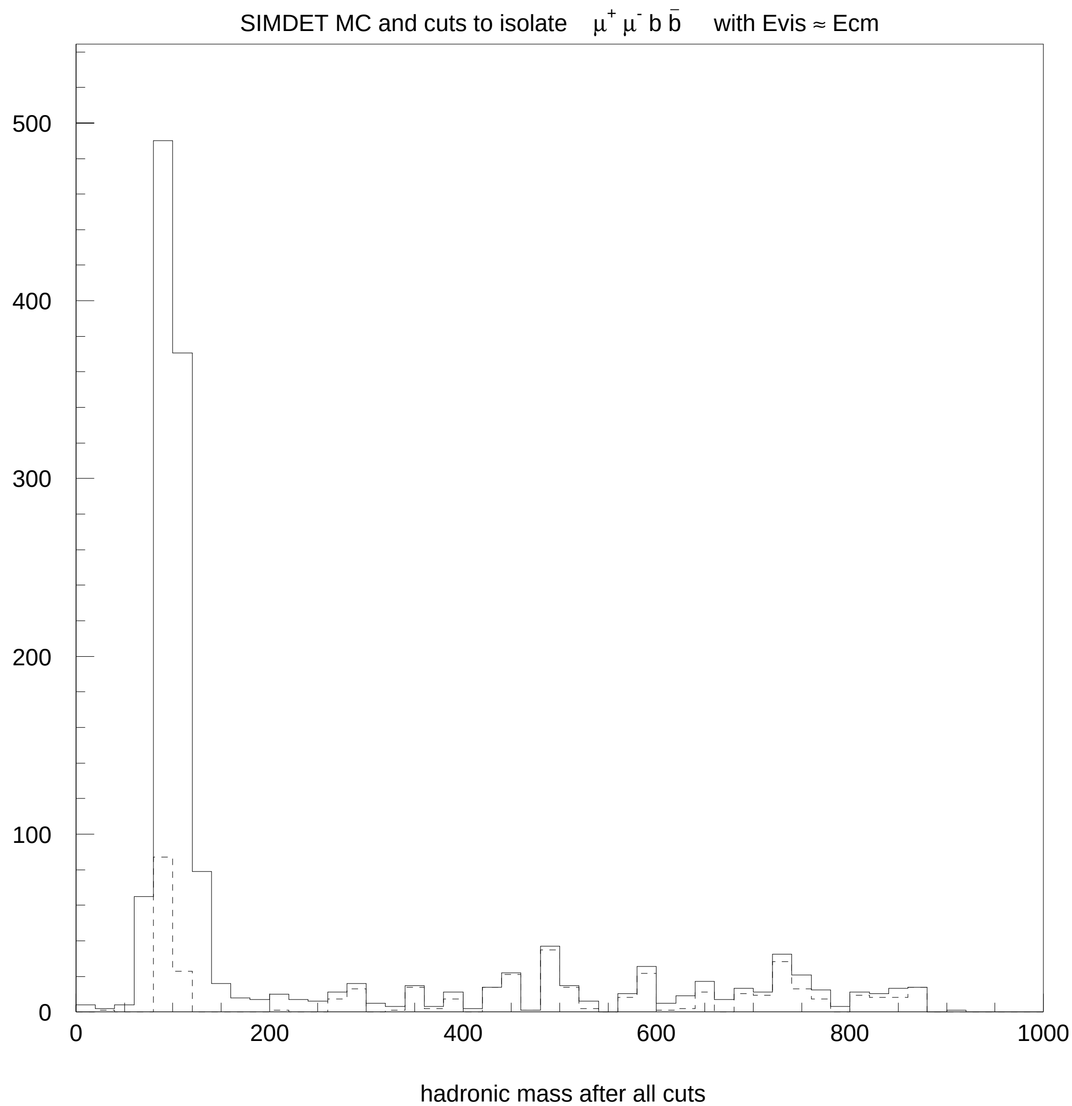
Use the following cuts:

- **Require $E_{vis} > 850$ GeV**
- **Require $|\cos\theta_{\text{thrust}}| < 0.8$**
- **Require two isolated muons with invariant mass between 82 and 100 GeV.**
- **Remove isolated muons and require that remaining system contain at least 6 charged tracks with momentum greater than 2 GeV and impact parameter greater than 3σ.**

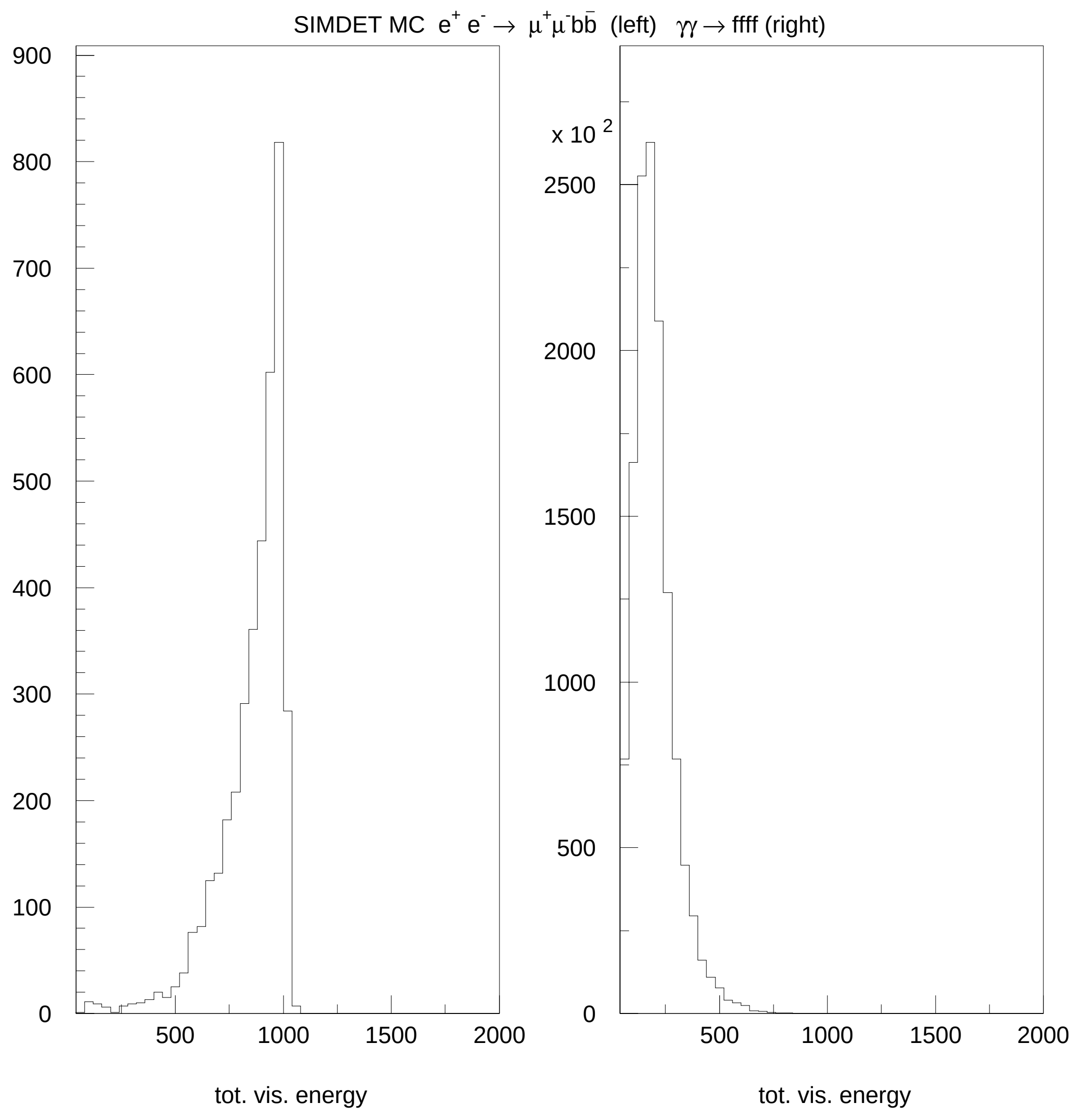
Solid histogram for ALL events, dashed for all events that are NOT from $e^+e^- \longrightarrow \mu^+\mu^- b\bar{b}$.



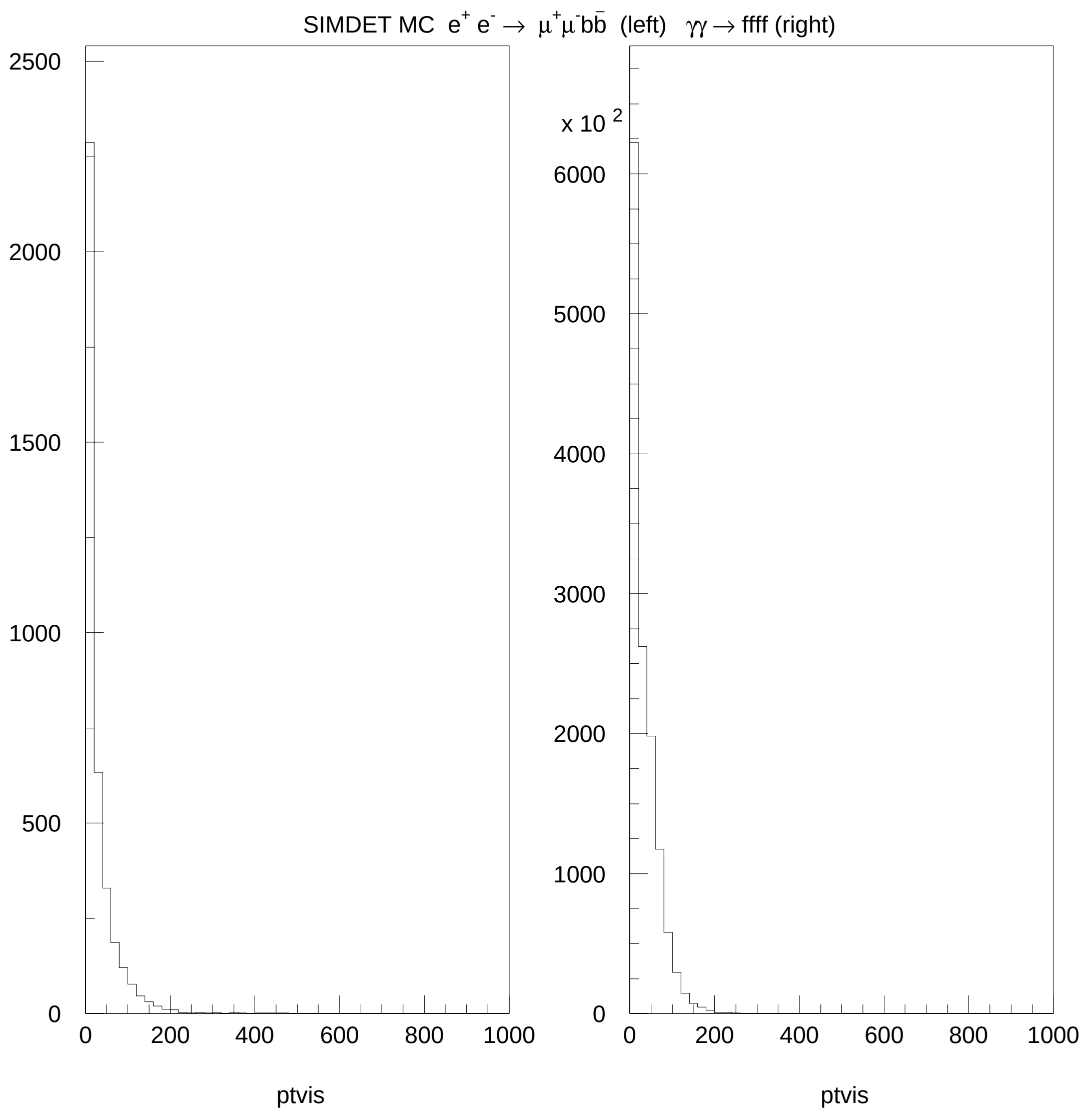
Solid histogram for ALL events, dashed for all events that are NOT from $e^+e^- \rightarrow \mu^+\mu^- b\bar{b}$.



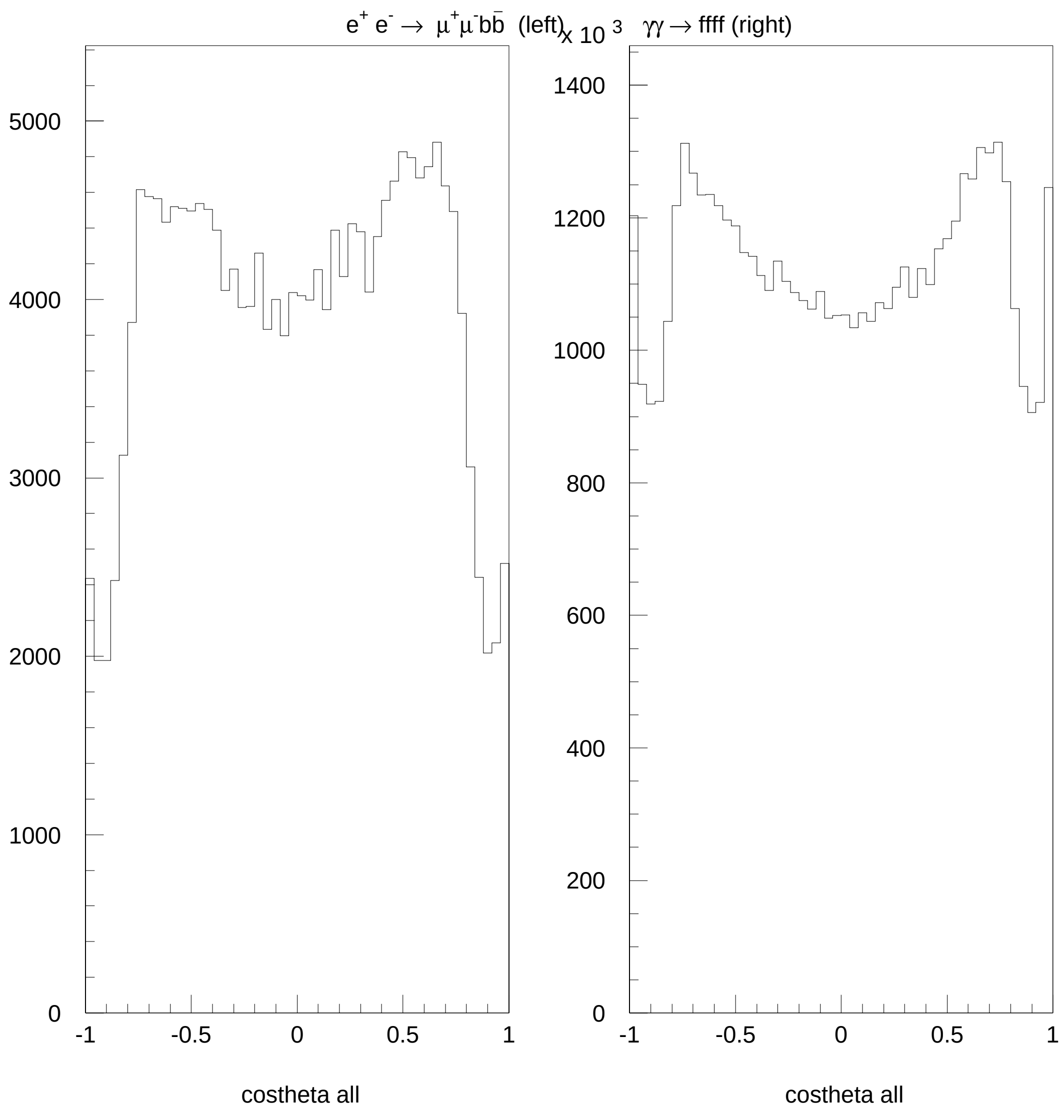
Total visible energy for $e^+e^- \to \mu^+\mu^- b\bar{b}$ (left) and $\gamma\gamma \to f\bar{f}f\bar{f}$ (right).



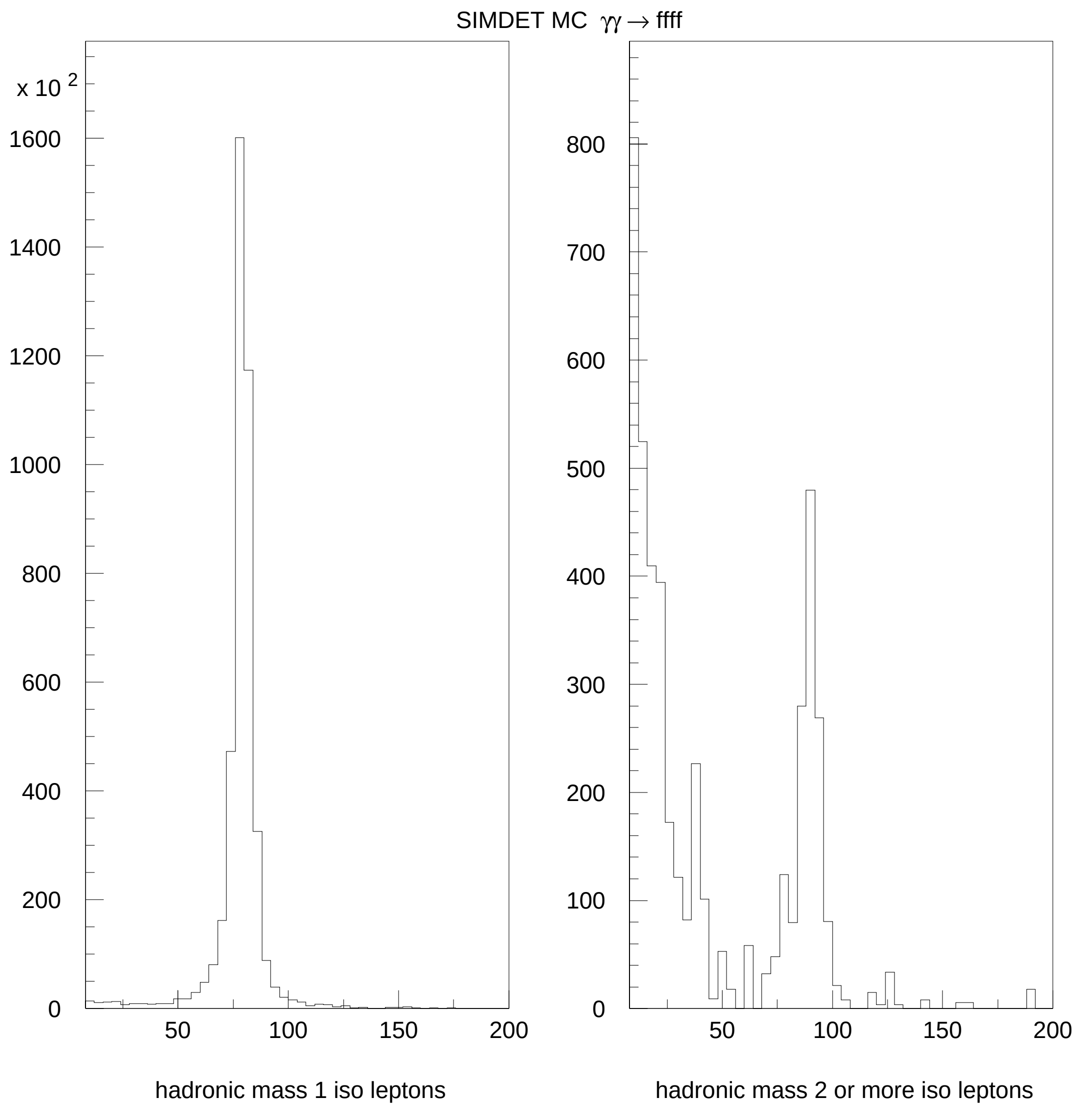
Visible P_T for $e^+e^- \to \mu^+\mu^- b\bar{b}$ (left) and $\gamma\gamma \to f\bar{f}f\bar{f}$ (right).



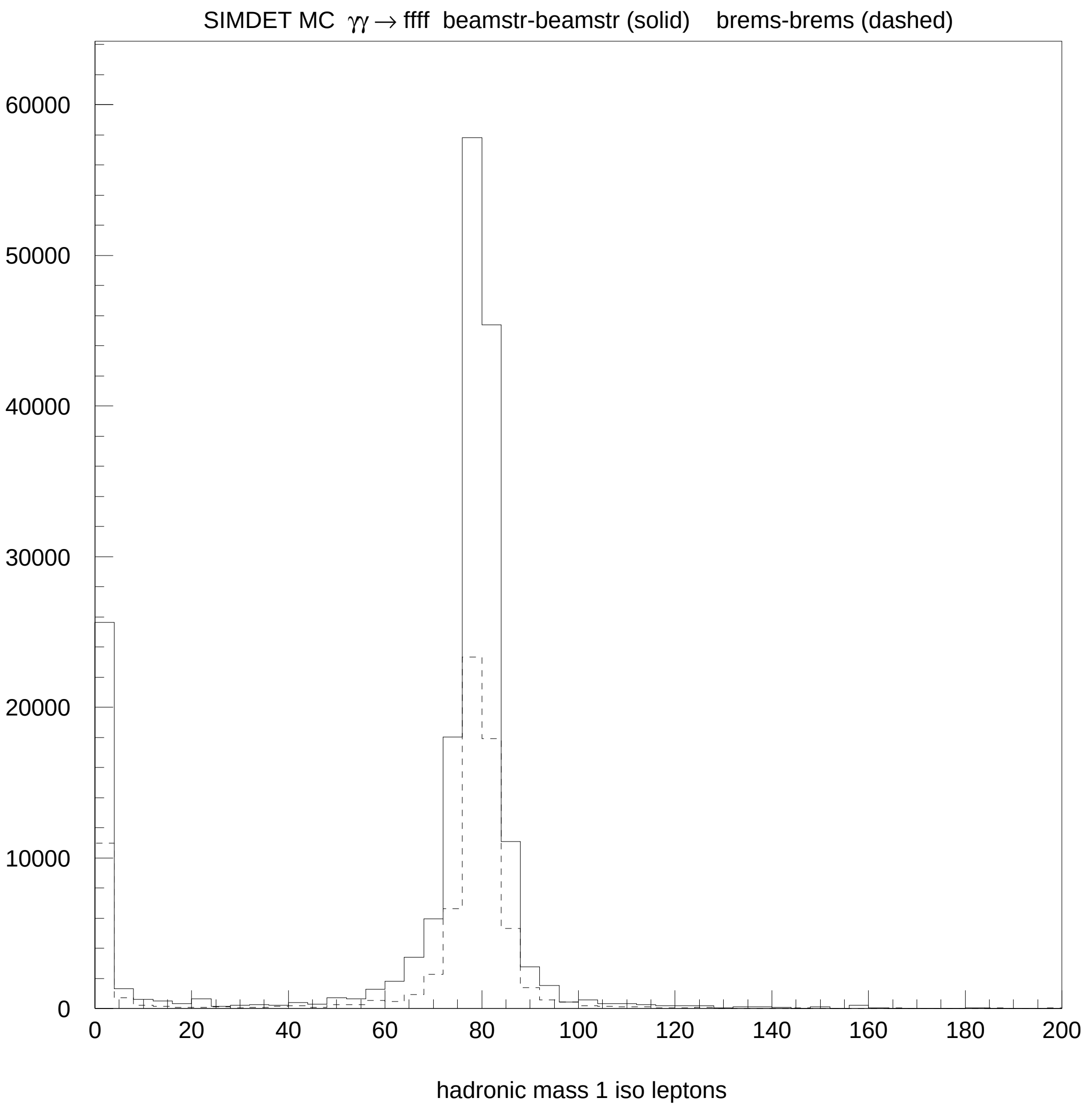
$\cos\theta$ for all charged and neutral tracks for $e^+e^- \longrightarrow \mu^+\mu^- b\bar{b}$ (left) and $\gamma\gamma \longrightarrow f\bar{f}f\bar{f}$ (right).



Mass of hadronic system for $\gamma\gamma \to f\overline{f}f\overline{f}$ for 1 isolated lepton (left) and 2 isolated leptons (right).



Breakdown of $\gamma\gamma \to f\bar{f}f\bar{f}$ according to beamstrahlung on beamstrahlung (solid) and bremsstrahlung on bremsstrahlung (dashed).



$e^+e^- \to \nu\overline{\nu}b\overline{b}$ **only on the left; ALL events on the right. Only requirements are that** $|\cos\theta_{\rm thrust}| < 0.8$ **and that there are no isolated leptons. B-tagging cuts and** E_{vis} **cuts will be needed to isolate Higgs signal on the left.**