CMS Internal Note

The content of this note is intended for CMS internal use and distribution only

23 April 2007

HEEP (High Energy Electrons and Photons) strategy toward first CMS data

J. Brooke, D. Evans, C. Hill, J. Jackson[a], D. Newbold[a]

H.H. Wills Physics Laboratory, University of Bristol, UK

O. Charaf, B. Clerbaux, G. De Lentdecker, Sh. Elgammal, P. Marage, M. Mozer, P. Vanlaer

IIHE (ULB-VUB), Université Libre de Bruxelles, Brussels, Belgium

V. Litvin, T. Orimoto

California Institute of Technology, Pasadena, USA

E. Carrera, Y. Gershtein

Florida State University, Florida, USA

A. Karneyeu[b][c], Ph. Miné, P. Nedelec[b]

LLR, Laboratoire Leprince-Ringuet, Ecole Polytechnique, Palaiseau, France

J. Cole, B. Kennedy, C. Shepherd-Themistocleous, R. Walton[d], S. Worm

RAL, Rutherford Appleton laboratory, Oxon, UK

P. Meridiani, A. Palma, R. Paramatti, C. Rovelli

Università degli studi di Roma ”La Sapienza”, Roma-1, Italy

N. Godinovic, I. Puljak, D. Polic

Technical University of Split - FESB, Split, Croatia

Z. Antunovic

University of Split - PMF, Split, Croatia

---

[a] also at RAL, Oxon, UK

[b] also at LAPP, Annecy, France

[c] on leave from INP, Minsk, Belarus

[d] also at Bristol University, UK

Abstract

This note proposes a strategy to prepare the analysis of high energy electron and photon pairs (HEEP) in the first CMS data, with two main goals: (i) the discovery of possible heavy resonances beyond the Standard Model, and (ii) the monitoring of the detector response, through the analysis of EW processes: e^+e^- pair production from the Drell-Yan process, di-photon production from $q\bar{q}$ annihilation and gluon fusion, and Zγ production. The emphasis will be on the high mass tail with $200 < M < 800~\mathrm{GeV/c^2}$ and, for the Drell-Yan process, also on the Z peak region.

A specific data stream for HEEP searches, collecting with very large efficiency electrons and photons with high transverse momentum, should be studied and will likely need to be integrated into the CMS trigger path. It will include – with limited bandwidth – both signal and control samples, in order to allow for the checking and monitoring of the selected data, to be performed mostly using the data themselves.

This note is an attempt to provide a list of tasks needed to define the data sets and to prepare for physics analyses. The objective is to be ready for the publication, based on a luminosity of $0.1~\mathrm{fb}^{-1}$, of a paper on “Search for massive resonance production decaying into an electron or a photon pair”.

1 Introduction

High mass resonances decaying into lepton or photon pairs provide some of the most important discovery potentials beyond the Standard Model (SM) at the LHC. They are predicted in numerous models (gravitons and gauge bosons in extra dimension models, new Z bosons in supersymmetric and GUT models, etc.) [1].

The search for pairs of electrons or photons (electromagnetic, "em" objects) with high p_t (transverse momentum with respect to the beam direction) is thus one of the hottest topics for CMS, from the very beginning of the data taking at 14 TeV.

From the LHC start up, the HEEP group (High Energy Electrons and Photons) [2] will be involved in the day-to day follow-up of data taking, with two main goals:

- the fast discovery of possible high mass resonances beyond the Standard Model, decaying into electron or photon pairs, with typically $M > 800~\mathrm{GeV/c^2}$, as lower mass resonances are already excluded at the Tevatron. This implies the progressive build-up of a resonance signal (in the ee and $\gamma\gamma$ channels) and the characterisation of its significance, including the understanding of the SM backgrounds in the mass spectrum (the irreducible EW backgrounds and the QCD jet backgrounds). The distortion of the expected mass spectra, in particular of the e^+e^- Drell-Yan spectrum, is another manifestation of new physics, which will require a detailed understanding of systematic effects; it might be difficult to achieve in early data taking;
- the study of the detector response to electrons and photons (trigger and detector efficiencies, particle identification, backgrounds). This could be performed in two complementary ways.
  (i) The detector response to electrons will be studied in the Z peak region and extrapolated to higher masses. This will allow collecting large di-electron statistics, but will suffer from the uncertainties due to the use of Monte Carlo simulations for the extrapolation.
  (ii) It will also be necessary to demonstrate that we understand the detector response to high energy electromagnetic objects, typically with $p_t > 80~\mathrm{GeV/c}$, using the data themselves. This will be monitored through the study of the high mass tail ($200 < M_{ee} < 800~\mathrm{GeV/c^2}$) of known electroweak (EW) processes: the Drell-Yan process for e^+e^- pair production, $q\bar{q}$ annihilation and gluon fusion for di-photon production, and Zγ production. This second method has the advantage to rely on data only but suffers from limited statistics.

Table 1 presents the production predictions for a SSM massive Z' boson and for a Randall-Sundrum (RS) extra-dimension graviton G (with coupling $c = 0.1$), with masses of 1.0, 1.5, 2.0 and 2.5 $\mathrm{TeV/c^2}$ [1] $^{1)}$. The expected numbers of events in the e^+e^- and (for G only) $\gamma\gamma$ channels are given for a luminosity of $0.1~\mathrm{fb^{-1}}$ and both decay particles in the tracker + ECAL acceptance ($|\eta| < 2.4$). A significant discovery potential is found for these channels, already with limited luminosity. The irreducible backgrounds (Drell-Yan production → e^+e^- and electroweak processes → $\gamma\gamma$) in the corresponding mass windows are also given. They are modest compared to the signals.

Table 2 presents the Drell-Yan cross section for $M_{ee} >$ 40, 120, 200 and 500 $\mathrm{GeV/c^2}$ and the expected numbers of events for several cuts on the electron transverse momenta, as defined with respect to the beam direction. The cross section at the Z peak is huge. With a minimum p_t of 12 $\mathrm{GeV/c}$ for both electrons as presently required at the L1 trigger level (see [4], p. 574, Table E.9), this corresponds to some 75,000 Drell-Yan events with both electrons having $|\eta| < 2.4$ for an integrated luminosity of $0.1~\mathrm{fb^{-1}}$. Practically, selection cuts applied at trigger and reconstruction level will slightly reduce this number. For the direct study of the detector response at large electron p_t, 140 Drell-Yan events will be recorded with $M_{ee} > 200~\mathrm{GeV/c^2}$. A handful of events will have a mass larger than $500~\mathrm{GeV/c^2}$.

The cross section for Drell-Yan production in the Z peak region will thus not be statistically limited with the first $0.1~\mathrm{fb^{-1}}$ of data, and the cross section measurement for $M_{ee} > 200~\mathrm{GeV/c^2}$ will be achieved with a statistical error of the order of 10%. After $1~\mathrm{fb^{-1}}$ has been accumulated, the analysis will be pursued with improved statistical precision, with a reach of higher masses and consideration of several mass bins.

Similarly, Table 3 presents the EW di-photon cross sections for $M_{\gamma\gamma} >$ 40, 120, 200 and 500 $\mathrm{GeV/c^2}$, and Table 4 presents the cross section for Zγ production, with an additional cut on the minimum transverse $\hat{p}_t$ in the hard process rest frame of 60 GeV/c. Note that the p_t cuts in Table 4 are applied on both leptons from the Z, as well as on the hard photon.

$^{1)}$ These numbers, as those of Tables 2 to 4, are obtained using the PYTHIA simulation program [3] at the generator level, with CTEQ5L parton distribution functions.

e^+e^- channel				
Model mass (GeV/c^2)	$M = 1000$	$M = 1500$	$M = 2000$	$M = 2500$
SSM Z'				
$\sigma \cdot$ BR (fb)	458	80	20	5.8
nb. events for 0.1 fb^{-1}	46	8.0	2.0	0.58
AND 2 electrons with $\|\eta\| < 2.4$	38	7.0	1.8	0.54
RS G ($c = 0.1$)				
$\sigma \cdot$ BR (fb)	660	76	14	3.5
nb. events for 0.1 fb^{-1}	66	7.6	1.4	0.35
AND 2 electrons with $\|\eta\| < 2.4$	62	7.2	1.3	0.32
irred. $q\bar{q} \rightarrow e^+e^-$ bg (GeV/c^2)	$M > 600$	$M > 1100$	$M > 1600$	$M > 2100$
cross section (fb)	50	4.4	0.76	0.18
nb. events for 0.1 fb^{-1}	5.0	0.4	0.08	0.02
AND 2 electrons with $\|\eta\| < 2.4$	3.9	0.4	0.07	0.02

$\gamma\gamma$ channel				
Model mass (GeV/c^2)	$M = 1000$	$M = 1500$	$M = 2000$	$M = 2500$
RS G ($c = 0.1$)				
$\sigma \cdot$ BR (fb)	1330	151	27.9	7.0
nb. events for 0.1 fb^{-1}	133	15.1	2.8	0.70
AND 2 photons with $\|\eta\| < 2.4$	113	13.5	2.6	0.66
irred. $q\bar{q} \rightarrow \gamma\gamma$ bg (GeV/c^2)	$M > 600$	$M > 1100$	$M > 1600$	$M > 2100$
cross section (fb)	224	29	6.6	2.0
nb. events for 0.1 fb^{-1}	22.4	2.9	0.7	0.2
AND 2 photons with $\|\eta\| < 2.4$	3.8	0.4	0.09	0.02
irred. $gg \rightarrow \gamma\gamma$ bg (GeV/c^2)	$M > 600$	$M > 1100$	$M > 1600$	$M > 2100$
cross section (fb)	5.5	0.21	0.021	0.0033
nb. events for 0.1 fb^{-1}	0.55	0.021	-	-
AND 2 photons with $\|\eta\| < 2.4$	0.22	0.009	-	-

Table 1: High mass e^+e^- and $\gamma\gamma$ resonance production and backgrounds: cross sections times branching ratios and numbers of expected decay particle pairs in the detector acceptance for a luminosity of 0.1 fb^{-1}, for a SSM massive Z' boson (decaying into an e^+e^- pair) and a Randall-Sundrum (RS) extra-dimension graviton G with coupling $c = 0.1$ (decaying into an e^+e^- or a $\gamma\gamma$ pair), with masses of 1.0, 1.5, 2.0 and 2.5 TeV/c^2; irreducible backgrounds: Drell-Yan $\rightarrow e^+e^-$ and EW di-photon production ($q\bar{q} \rightarrow \gamma\gamma$ and $gg \rightarrow \gamma\gamma$), in the corresponding mass windows.

Drell-Yan events (GeV/c^2)	$M > 40$	$M > 120$	$M > 200$	$M > 500$
cross section (fb)	1795 10^3	18,827	2524	99
nb. events for 0.1 fb^{-1}	179,500	1,883	252	9.9
AND 2 electrons with $\|\eta\| < 2.4$	77,773	916	144	7.2
AND $p_t(1,2) > 12$ GeV/c	75,061	912	144	7.2
AND $p_t(1,2) > 40$ GeV/c	20,140	747	138	7.2
AND $p_t(1,2) > 80$ GeV/c	182	142	105	7.1

Table 2: Cross section and numbers of expected Drell-Yan events with $M_{ee} > 40, 120, 200$ and 500 GeV/c^2, for a luminosity of 0.1 fb^{-1} and various electron p_t cuts.

$q\bar{q} \rightarrow \gamma\gamma$ events (GeV/c^2)	$M > 40$	$M > 120$	$M > 200$	$M > 500$
cross section (fb)	306,900	21,620	5546	398
nb. events for 0.1 fb^{-1}	30,690	2,162	554	40
AND 2 photons with $\lvert\eta\rvert < 2.4$	5807	384	98	6.7
AND $p_t(1,2) > 40$ GeV/c	480	278	88	6.7
AND $p_t(1,2) > 80$ GeV/c	81	73	60	6.4

$gg \rightarrow \gamma\gamma$ events (GeV/c^2)	$M > 40$	$M > 120$	$M > 200$	$M > 500$
cross section (fb)	331,000	6245	818	13.5
nb. events for 0.1 fb^{-1}	33,100	625	82	1.3
AND 2 photons with $\lvert\eta\rvert < 2.4$	11,540	220	30	0.53
AND $p_t(1,2) > 40$ GeV/c	309	141	26	0.53
AND $p_t(1,2) > 80$ GeV/c	13	21	15	0.50

Table 3: Cross section and numbers of expected EW di-photon events with $M_{\gamma\gamma} > 40, 120, 200$ and 500 GeV/c^2, for a luminosity of 0.1 fb^{-1} and various electron p_t cuts.

Zγ events (GeV/c^2)	
cross section (fb)	385
nb. events for 0.1 fb^{-1}	38.5
AND l, γ with $\lvert\eta\rvert < 2.4$	18.3
AND $p_t(l,\gamma) > 40$ GeV/c	5.9
AND $p_t(l,\gamma) > 80$ GeV/c	0.7

Table 4: Cross section and numbers of expected Zγ events (in the dielectron and dimuon Z decay channels), for a luminosity of 0.1 fb^{-1} and various lepton and photon p_t cuts. The events were generated with a cut on the minimum transverse $\hat{p_t}$ in the hard process rest frame of 60 GeV/c.

This note is an attempt to provide a list of options and tasks needed to prepare physics analyses and to be ready for the fast publication, based on a luminosity of 0.1 fb^{-1}, of a paper on "Search for massive resonance production decaying into an electron or a photon pair". It is organised as follows.

Section 2 presents studies of the L1 and HLT triggers. It is shown in Section 2.2 that the standard HLT selection criteria for electrons and photons induce very significant event losses at high mass, i.e. large p_t of the decay particles.

As discussed in Section 3, a specific trigger path for HEEP searches will likely need to be defined, to collect with very high efficiency electrons and photons with a large p_t. It should include
- data samples directly relevant for the high mass resonance search and for studies of the EW processes (Section 3.1);
- detector oriented control samples, needed to perform checks of the selected data quality, in particular for trigger efficiencies and particle identification (Section 3.2).

The data quality monitoring of the HEEP stream, which should be performed quasi-online, is discussed in Section 3.3.

The search for new resonances using di-em objects is discussed in Section 4.1. The measurement of the Drell-Yan cross section is presented in Section 4.2. A first attempt to discuss efficiency measurements and background estimates using data is performed in Sections 4.2.2 and 4.2.3. Similarly the measurements of di-photon EW processes at high mass and of Zγ production are very briefly introduced in Section 4.3. Systematic uncertainties are discussed in Section 4.4.

The note finally contains remarks concerning the pilot run at 900 GeV (Section 5).

All relevant studies will be performed in collaboration with the e/γ Physics Object Group, with the Electroweak and Online Selection Physics Analysis Groups, and in connexion with other contributors to the High Energy Pairs study group of the SUSY/BSM PAG.

2 Present L1 and HLT triggers

Specific studies of the L1 and HLT trigger performance for high energy electrons and photons, performed within the HEEP group, are summarised.

2.1 L1 and saturation effects

Effects on the L1 trigger of the ECAL electronics saturation at high energy have been investigated (see note in preparation [5]).

It is known that the ECAL electronics (MGPA + ADC on VFE cards) saturates at approximately 1500 GeV in the barrel and 3000 GeV in the endcaps. It has been feared that this could affect the timing corresponding to the maximum of the signal collection, and thus hamper the event read out. However, it is found that saturation effects do not destroy the ability of the TPG to identify the correct bunch crossing for read out.

On the other hand, a hysteresis effect over 5 signal samples forbids the use of the falling edge of the signal to reconstruct the full amplitude, and software corrections to saturation will thus remain necessary (cf. [6]).

In coordination with the ECAL DPG, the necessary actions are being taken in order to ensure that saturating events are correctly registered, and that saturation effects are properly described in the simulations.

The investigation of the L1 trigger efficiencies for high p_t electrons and photons is one of the most important tasks of the HEEP group. It has to be performed with the L1 emulator, within the CMSSW framework.

2.2 HLT efficiencies at high p_t

The present HLT trigger definition for the di-electron stream has been tuned to optimise Higgs boson discovery in the $H \to ZZ \to 4\ e$ channel, with electron $p_t \leq 40$ GeV/c, and with acceptable QCD background rates. It requires [7]:
- a transverse momentum for each electron candidate $p_t > 12\ \mathrm{GeV/c}$;
- less than 9 GeV deposit in the HCAL behind the ECAL supercluster (HCAL isolation cut);
- matching with the pixel detector,
- a track isolation criterium.

	DY $> 200\ \mathrm{GeV/c^2}$	DY $> 500\ \mathrm{GeV/c^2}$	Z' $2\ \mathrm{TeV/c^2}$	RS G $2\ \mathrm{TeV/c^2}$
Di-electron HLT stream				
$p_t > 12\ \mathrm{GeV/c}$	98.5 %	98.7 %	99.3 %	99.5 %
HCAL < 9 GeV	94.6 %	90.6 %	68.5 %	69.2 %
pixel matching	85.3 %	88.8 %	83.5 %	80.0 %
track isolation	87.7 %	88.1 %	87.4 %	86.3 %
total efficiency	69.6 %	69.9 %	49.6 %	47.5 %
Single electron HLT stream				
$p_t > 26\ \mathrm{GeV/c}$	98.2 %	99.9 %	99.9 %	99.9 %
HCAL < 3 GeV	96.7 %	96.1 %	54.4 %	48.5 %
pixel matching	95.8 %	97.2 %	92.4 %	88.4 %
track isolation	95.0 %	95.9 %	93.8 %	94.7 %
E/p cut	87.6 %	81.6 %	77.9 %	77.4 %
total efficiency	75.7 %	73.0 %	36.7 %	31.4 %

Table 5: Effect of the HLT cuts on the reconstruction efficiencies of Drell-Yan states (with $M > 200$ and $M > 500\ \mathrm{GeV/c^2}$) and of Z' bosons and RS gravitons ($c = 0.1$) with masses $M = 2\ \mathrm{TeV/c^2}$.

Table 5 presents the detection efficiencies corresponding to these requirements (as applied successively) and the cumulative efficiency, for e^+e^- events with both electrons emitted in the tracker + ECAL region with $|\eta| < 2.5$ [2].

---

[2] In view of the absence of the L1 emulator at the time of this analysis, the efficiencies reported here were obtained by comparing the number of HLT selected events to the number of generated events, with no simulation of the L1 trigger effect. They are thus describing in fact the convolution of the L1 and HLT triggers.

The considered channels are Drell-Yan production with masses $M_{ee} > 200$ and > 500 GeV/c^2, and the production of SSM Z' bosons and Randall-Sundrum gravitons with masses of 2 TeV/c^2.

For comparison, the table also presents the efficiencies when the single-electron stream requirements are fulfilled by at least one of the electrons. The latter case implies different numerical values for the cuts (see table), and in addition a matching between the electron momentum p measured by the tracker and the energy E measured by the ECAL ($E/p < 1.5$ in the barrel, and $E/p < 2.45$ in the endcaps).

The overall efficiencies are low for high mass electron pairs. The HCAL cuts at fixed values (either 3 or 9 GeV) are inadequate at high energy because of the longitudinal shower leakage [6]. The E/p matching requirement for the single-electron stream is affected by the error on the momentum measurement for rather straight tracks.

	RS G 2 TeV/c^2
Di-photon HLT stream	
$p_t(1) > 30$ GeV/c, $p_t(2) > 20$ GeV/c	98.9 %
ECAL isolation cut ($<$ 1.5 GeV)	58.8 %
HCAL $< 8(6)$ GeV	43.1 %
tracker isolation cut	94.2 %
total efficiency	23.6 %
Single photon HLT stream	
$p_t > 80$ GeV/c	99.9 %
ECAL isolation cut ($<$ 1.5 GeV)	86.8 %
HCAL $< 6(4)$ GeV	30.4 %
tracker isolation cut	87.2 %
total efficiency	23.0 %

Table 6: Effect of the HLT cuts on the reconstruction efficiency of a Randall-Sundrum graviton decaying into two photons ($M = 2$ TeV/c^2, $c = 0.1$); the quoted HCAL isolation cuts are, respectively, for the barrel and endcap regions.

Table 6 similarly presents the reconstruction efficiency for a Randall-Sundrum graviton decaying into two photons, using the present HLT criteria tuned for the $H \rightarrow \gamma\gamma$ channel and the corresponding background rates. The low ECAL and HCAL isolation cuts induce very large losses.

3 Trigger path for HEEP searches

For the search of heavy resonances beyond the Standard Model, the low efficiencies at large mass apparent in Tables 5 and 6 are problematic in view of the small statistics to be expected, and also because of the large uncertainty which will affect such low efficiencies, especially at the beginning of data taking.

Specific trigger paths are thus needed for the HEEP searches, in order
- to collect the high p_t di-electron and di-photon signals with very large efficiencies;
- to permit, using control samples, the measurement of efficiencies and the estimate of backgrounds.

The stream criteria, defined in collaboration with the e/γ POG and the Online Selection Group, will result from a compromise between a high efficiency in the selection of useful data and of a limited bandwidth, with limited background rates. The rates will be studied using simulations of QCD jets and of the EW processes discussed in this note; for the di-photon channel, prompt photon emission will also be considered.

The total HLT rate is expected to be limited to some 100 Hz. Of this, a bandwidth of about 20 % could be available for electromagnetic objects, including Z and W production, data for detector calibration, Higgs and SUSY/BSM searches. The data necessary for HEEP analyses will be at least partly included in other em streams, a most striking example being Z production. Of the full data collection, 5 to 10 %, i.e. around 10 Hz, is attributed to the Express Data Stream [8]. The additional HEEP contribution will be a very small fraction of this bandwidth. It should be noted that the true Drell-Yan bandwidth above 40 GeV/c^2 is 0.15 Hz in the detector acceptance for a luminosity of 2 10^{32} cm^{-2}s^{-1} at start-up, and is negligible above 120 GeV/c^2 (2 10^{-3} Hz).

To adapt quickly to the available bandwidth in the presence of the data, several schemes will be defined, corresponding e.g. to 1.0, 0.5, 0.2 or 0.1 of the a priori expected bandwidth. They will certainly have to evolve with luminosity.

3.1 Signal samples

One possibility is to select a signal data sample made of two em objects (**di-em sample**), or possibly two electrons or two photons separately. An alternative is to define for the signal collection a **single em** (or single electron) stream with (possibly p_t depending) isolation cuts. In any case, the HEEP stream should collect high p_t electron and photon pairs with very high efficiency.

Following the discussion of the effects of the present HLT requirements in Section 2.2, it appears desirable, in order to achieve high efficiencies, to relax the HCAL isolation and E/p cuts.

In view of the uncertainties affecting the detector response at the beginning of data taking, it is probably desirable to relax track requirements (requirements of a track–cluster association and of E/p matching for electron identification; requirement of no track for photon identification). This will allow to treat electron and photon candidates on the same footing, and thus remove dependences on tracking efficiency for electrons and on conversion rate for photons. Studies with also be performed on the requirement of (possibly loose) track association, to be used either already at start-up if required by the background rates, or at a later stage when the detector is better understood.

Because of the uncertainties affecting the physics of the underlying event, it is probably also desirable to keep isolation criteria around the electron or photon candidates as loose as possible.

For very loose cuts on the di-em objects (two ECAL clusters, each with $p_t > 80$ GeV/c and with HCAL / ECAL (H/E) energy deposits < 0.05 or HCAl energy smaller than 9 GeV), preliminary studies indicate that a very high Drell-Yan signal efficiency is achieved with very low QCD rates : $\sim$0.1 Hz for an instantaneous luminosity at start-up of 2 10^{32} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. As the p_t threshold is decreased, the rates increase to $\sim$2 Hz for $p_t > 50$ GeV/c, $\sim$130 Hz for $p_t > 30$ GeV/c. The possibility will thus be studied of applying additional isolation cuts to decrease the rate in the low p_t region, relevant for the Z peak, and / or of keeping loose cuts with a prescale factor.

Similar cuts on a single em object in the large p_t region (one ECAL cluster with $p_t > 80$ GeV/c and with H/E energy deposits < 0.05) induce a 50 times larger QCD background rate, of $\sim$5 Hz. In this case, additional isolation cuts and / or track-cluster association will be compulsory in the lower p_t range, especially to cover the Z peak region with acceptable QCD background rates. A p_t dependence of these criteria is to be studied, in order to make sure that an unbiased sample is collected at high p_t, with very loose selection cuts.

The following checks and studies need to be performed:

- optimisation of the p_t thresholds, possibly with different values in case of two em objects (asymmetric p_t cuts);
- additional cuts on the azimuthal angle difference $\Delta\phi$ between the two em objects and / or on their invariant mass M;
- optimisation of the HCAL / ECAL isolation cut with a possible continuous p_t dependence;
- optimisation of the (p_t dependent) cut on the energy content E_t inside a given $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ cone around the em object
  - defined either in absolute or in relative terms
  - using ECAL and / or HCAL and / or tracker objects;
- study of the possible track association requirements for electrons;
- similarly, study of the opportunity of requiring no track association for photons.

Other optimisations are concerned e.g. with:

- the definition of robust ECAL clustering algorithms at start up for the HEEP data set, in order
  - to minimise losses, e.g. due to Bremsstrahlung effects on the ΔR isolation;
  - to take into account a possibly large ECAL noise.

  Studies are presently performed, indicating that 3×3 clusters might be optimal, and proposals for implementations in the trigger code are being prepared, in collaboration with the ECAL PDG, e/γ POG and OnSel groups (see note in preparation [9]);

- track–cluster linking algorithms for large p_t electrons with straight tracks, with possibly a specific use of pixels in collaboration with the e$/\gamma$ POG;
- electronic saturation effects [5];
- the possibility of triggering high p_t em objects using the HCAL endcap longitudinal segmentation only, in case the ECAL endcaps are not fully in place at start-up.

3.2 Control samples

Control samples will be needed to measure trigger and particle identification efficiencies, and to determine backgrounds under a possible high mass resonance and under the Drell-Yan and di-photon signals (see Sections 4.2.2 and 4.2.3).

Two regions will be studied, with masses below the discovery domain ($M > 800$ GeV/c^2) :
- the Z peak region;
- the high mass range with $200 < M < 800$ GeV/c^2.

For efficiency and background estimates, it will probably be needed to collect a single em sample, based on a trigger path requiring only one em object. This data sample may consist of events with loose isolation cuts and possibly a large prescale factor, and of samples of events with tighter cuts (including possible track-cluster association criteria) and smaller prescales. Cuts and prescale factors may be p_t dependent.

Different prescale factors need to be prepared, according to the actual rates at data taking.

3.3 Data quality monitoring for the HEEP stream

At the beginning of the data taking, trigger rates and efficiencies, detector response and particle identifications need to be followed closely. The quasi "on-line" follow-up of the HEEP stream will be achieved using real events, selected both from the signal and the control samples.

Histograms will be prepared in collaboration with the DQM Group, in order to catch rapidly with detector problems, understand trigger performance, and eventually set prescales if needed. These histograms will be updated and compared on a day-to day basis with expectations, both for the shapes of the distributions and for the absolute rates.

The DQM follow-up will concern:

- trigger monitoring:
  - L1 and HLT bits, rates and disk space occupancy;
  - hot channels should be quickly identified and discarded from the trigger;
- checking of the detector response:
  - general detector features: ϕ distributions, $\eta > 0$ and $\eta < 0$ comparison, e^+ and e^- symmetry in the pair cms;
  - ECAL control plots : cluster shapes (S1, S9, S24, S25, X - Y variables, saturation), clustering algorithms, consistency between ECAL barrel and endcaps, energy deposit in the preshower;
  - H/E distribution versus ECAL energy;
  - energy distribution within the ΔR cone.

In addition, other streams, in particular the dijet stream, will be checked for possible losses of interesting events from our main stream.

4 Physics analyses

The physics analyses at the beginning of data taking will be devoted to the search for massive resonance production – which is the physics objective –, and to the measurement of the reference EW processes – intended at checking and demonstrating our understanding of the CMS detector response to high energy em objects.

4.1 Search for heavy resonances

The basic tool for the discovery of heavy Z' and graviton resonances beyond the Standard Model is the recording of high mass di-electron and di-photon pairs, $M > 1000 \text{ GeV/c}^2$, and the study of their characteristics for model discrimination.

In addition to the building-up of heavy resonances, modifications of the SM Drell-Yan spectrum around 1 TeV can also indicate the presence of new physics, in particular due to extra spatial dimensions. Useful indications or limits may be already accessible for 1 fb^{-1} [1]. This requires a detailed understanding of the detector response, in view of the steeply falling Drell-Yan spectrum and of energy smearing effects.

The selection efficiency of high p_t electrons and photons (typically $p_t > 300 \text{ GeV/c}$) from the decay of a heavy resonance should be as high as possible, close to 100%, in view of the possibly low statistics in the signal. In addition, low selection efficiencies would be affected by large uncertainties, due to the uncertainties in the extrapolations from the Z region, and to the small statistics available to determine efficiencies at large p_t using the data themselves (see discussion in Section 4.2.2).

Hence the identification requirements must be loose. At high energy, a significant H/E longitudinal leakage is indeed expected, the track parameter measurements (momentum, charge) are poor, and the ΔR isolation criteria must be handled with care, in view of the unknowns concerning the underlying event in the presence of a very heavy resonance.

As shown in the TDR-II [1], the instrumental backgrounds due to fake electron or photon pairs from QCD jets are expected to be quite small at high energy. This must be confirmed within the data, in the context of a poorly calibrated and aligned detector at start-up. The irreducible EW backgrounds have small cross sections at high mass, as shown in Table 1.

The background under the resonant signal will therefore be estimated mainly using the data in the sidebands. As a cross check, SM model background expectations will also be estimated using simulations.

In case evidence is found for resonance production, the following will need to be performed:
- evaluation of the statistical significance of the signal;
- search for signal confirmation in other decay channels, both in the HEEP stream (ee and $\gamma\gamma$ signals) and in other streams ($\mu\mu$, di-jets);
- determination of the production and decay couplings;
- investigation of the resonance spin using angular distributions.

4.2 Measurement of Drell-Yan production

The strategy for HEEP searches relies on the measurement of Drell-Yan production. It is of primary importance for the search of a BSM signal in the di-electron and di-photon channels to demonstrate that the detector response is under control, both at the Z peak and in the high mass region with $M > 200 \text{ GeV/c}^2$.

This will be performed through the measurement of the total and differential Drell-Yan cross sections, and the comparison with the SM predictions obtained from theory and from Tevatron measurements. Conversely, discrepancies between Drell-Yan measurements and SM predictions in the region with $M < 800 \text{ GeV/c}^2$ would point to uncontrolled backgrounds, trigger losses or shortcomings in high energy electron or photon identification or reconstruction. Although the precision extraction of SM parameters from Drell-Yan measurements is not a specific objective of the HEEP group, the understanding of the M_{ee} spectrum from the Z peak to the high mass tail will constitute one of the main checks of the HEEP data selection at the beginning of data taking. This will be performed in collaboration with the EW Group.

Two complementary ways of checking the detector response may be used, possibly jointly.

A first possibility, taking advantage of the large statistics, is to estimate the trigger, selection and identification efficiencies at the Z peak, to check their p_t dependence using the data, and to extrapolate to high p_t and high mass. The selection variables and the effects of the cuts, as well as the detector response, thus need to be very well understood and modeled, preferably with weak p_t dependence.

On the other hand, since the HEEP searches are concerned with a p_t domain where the detector is not calibrated, and far away from the p_t range relevant for the Z peak, Monte Carlo simulations need to be considered with care, especially at start-up. It would thus be important to also estimate efficiencies and backgrounds for the different sets of selection criteria, using the data themselves at (relatively) high p_t.

4.2.1 Drell-Yan event selection criteria

The QCD backgrounds in the mass region of interest will be large. In order to decrease the background level well below the signal, additional selection cuts for electron identification will need to be applied.

The electron identification criteria for Drell-Yan event selection may include for example the following:
- cuts on the H/E deposits (possibly different for barrel and endcaps);
- ΔR isolation criteria, using tracks, ECAL and HCAL clusters;
- track–cluster association criteria;
- E/p matching;
- total charge 0 of the di-electron pair.

Before data taking, the cuts and criteria values will have to be determined using Monte Carlo simulations. Several sets of selection criteria will be prepared, which will allow
- to perform cross checks between sets of criteria;
- to react fast in the presence of the data, following the detector performance and the actual rates.

The selection criteria will evolve in view of the data themselves, in order to enhance signal efficiency and purity.

4.2.2 Trigger and electron identification efficiency measurements

As presented above, the trigger and electron identification efficiencies will be measured (with high statistics) at the Z peak, and then extrapolated to the high mass region, with inherent uncertainties. In parallel, they will be measured at high mass from the data themselves, with however limited statistical precision.

At the Z peak, where the ratio of the Drell-Yan signal to the QCD background is favourable, the ”tag and probe” method is used to estimate the selection and identification criteria efficiencies, and to check the simulations. The overall Z production cross section will be compared to the Standard Model expectation.

In the high mass tail, where the data only are to be used, a **golden electron** sample will be defined with strict cuts on the electron selection, for example: HCAL energy deposit < 2 GeV (i.e. consistent with 0, but allowing for noise fluctuations); a well associated track, with a well defined charge ($> 3\ \sigma$); E/p cluster energy / track momentum matching (numerical values to be defined, for barrel and for endcap regions); track / cluster isolation criteria (to be defined); small missing E_t of the event, in order to discard W decay contributions...

The basic hypothesis is that, for events containing a well-identified (golden) electron, the only significant SM process is e^+e^- Drell-Yan production. The presence of a golden electron in an event would thus imply that a second electron, with opposite charge, has been produced in the same event.

- The HLT trigger efficiency for a di-em stream may be determined using the single em sample. If an event contains a golden electron and no other electron candidates, it may indeed be lost for one of the following reasons:
  - the second electron was emitted beyond the detector geometric acceptance (i.e. large $|\eta|$); the corresponding fraction of lost events can be calculated using the simulation of Drell-Yan production and detector geometry;
  - the second electron has been lost at the di-em trigger level.

  More specifically, some of the losses related to the H/E cut are due to the fact that the second electron was emitted in the direction of a crack between ECAL modules or supermodules or in the direction of the gap between the ECAL barrel and endcaps, and has consequently not passed the ECAL p_t threshold or has deposited a significant energy fraction in the HCAL. The corresponding efficiency estimate can be cross-checked using a technique which permits to recover part of such electrons using energy deposits in the HCAL (see note in preparation [10]).

- With events containing a golden electron and a second ECAL cluster with high p_t, the electron selection efficiency and its dependence on various identification criteria can be studied using the unbiased probe electron which is the companion to the golden one. The efficiency can be cross checked using different sets of identification criteria, in particular by requesting two golden electrons.

- The distributions of the identification variables for the electron candidates which do not pass the final selection criteria can be studied. This provides consistency checks between the distributions in these events and the estimated number of lost events.

- The track association efficiency (i.e. electrons appearing as γ candidates) and the wrong charge assignment may be measured using di-em events containing a golden electron, since following the above hypothesis the second em object should be an electron with the right charge.

4.2.3 Background estimates

The QCD background under the Drell-Yan signal needs to be studied, and in particular the mass and the selection criteria dependence of the signal / background ratio.

In the Z mass region, the resonant peak itself will provide handles for the background estimates.

In the high mass tail with $200 < M_{ee} < 800 \text{ GeV/c}^2$, background estimates may be provided by the study of the following:

- The false electron identification may be determined using the single em sample, by selecting events containing, in addition to the electron candidate, a single-jet with p_t above some given threshold, opposite back to back with the electron candidate, and with small missing E_t in order to discard W decay contributions. Such events are most probably dijet events with false electron identification.
  The ratio of the number of these events to the total number of dijet events, selected with the same p_t threshold and the same missing E_t, will allow to determine the probability of jet misidentification as an electron, and thus provide an estimate of the dijet background in the di-em sample.
- In the di-em sample, the fraction of QCD events with false identification of both electrons can be obtained from the following samples:
  - like-sign di-electron events (e^+e^+ and e^-e^-), passing the final selection criteria except for the total charge; after subtraction of the number of events with wrong charge assignment, the remaining number of events is due to two false electron identifications;
  - similarly for events with one identified electron and one photon candidate (i.e. with no associated track) satisfying all other selection criteria: after correction for track reconstruction inefficiency, the remaining number of events is due either to prompt photon events with a false electron, or to two false em identifications.
  From this, the probability of false identification and the number of false e^+e^- events in the Drell-Yan selected sample can be thus be extracted.
- The study of the sample of di-em events which do not pass the final Drell-Yan selection criteria will allow to determine, in addition to dijet events, which other topologies (e.g. multi-jets) lead to backgrounds. The same study can thus be made as for dijets.
- The features of the sample of events selected in the di-em stream but which do not pass the final Drell-Yan selection criteria will be compared to the SM Monte Carlo expectations. In spite of the fact that we are dealing here with tails of distributions, Monte Carlo simulations can provide useful guidance, especially at the beginning of data taking.

The sample of events with 1 electron candidate + 1 opposite jet as defined above will be used to test the purity of the golden electron sample. The assumption that golden electron are due to Drell-Yan event decays would imply that no such events should be detected. Conversely, the number of events with 1 golden electron + 1 opposite jet allows to determine the background contamination in the golden electron sample.

4.2.4 Comparison with SM predictions

The SM predictions for Drell-Yan cross section will be calculated using Monte Carlos simulations, including detector degradation and pile-up at start up.

Theoretical uncertainties (QCD and EW higher orders, effects of pdf uncertainties, underlying event, etc.) will be studied, in particular in connexion with the dedicated Les Houches effort [11]. They may however be expected to be smaller than the statistical and experimental uncertainties at the beginning of data taking.

All relevant variables will be systematically compared to predictions in control plots, including:
- for single electrons (separately for e^+ and e^-): p_t, p_z, η, ϕ, H/E, energy deposit within ΔR;
- for the e^+e^- pairs: invariant mass distribution, p_t, p_z, η, ϕ of the pair, e^+e^- asymmetry;

- for the overall structure of the event and the underlying event: numbers of observed tracks, clusters and jets; total energy deposit in the ECAL / HCAL / ECAL + HCAL / forward calorimeters; sphericity parameters (thrust, coplanarity, etc.) of the underlying event after removal of the 2 electrons.

4.3 Measurements of high mass di-photon production and of Zγ production

The analysis of the di-photon reference samples can be conducted in a way similar to the analysis of the Drell-Yan process discussed in detail in the previous Section, in the high mass tail ($M > 200$ GeV/c^2), as no Z peak may be used here. This will thus not be repeated here. A specific background channel is however prompt photon production (γ + jet), with a misidentified second photon.

For Zγ production, the signal should be clean if strong selection requirements are applied. These have to be defined, taking into account that the photon and the Z should have large p_t, and be emitted almost back to back.

4.4 Other systematic error estimates

The estimate of systematic uncertainties will require careful studies, both for the discovery of a BSM signal and for Drell-Yan measurements.

- In the case of the discovery of narrow resonances with mass above $1\ \mathrm{TeV/c^2}$, systematic effects in the signal extraction and the total cross section measurement are expected to be small. This is also true, although more difficult, in the case of broad peaks, in view of the expected limited background at high mass.
  On the other hand, the determination of the resonance mass will rely on the ECAL calorimeter calibration, for which no direct method exists at high energy.
  With large signal statistics, the measurement of asymmetries and the determination of the spin structure of the signal will imply the unsmearing of several effects, in particular related to the ignorance of the initial quark (antiquark) direction.
- The control of systematic uncertainties of instrumental origin will be demonstrated through the good understanding of the Drell-Yan spectrum, from the Z peak up to the high mass region:
  - the accuracy of the detector calibration at high energy (linearity of the response) will be estimated through the description of the Drell-Yan mass spectrum itself in the SM region as determined at the Tevatron (up to some $800\ \mathrm{GeV/c^2}$);
  - a kinematic method is under study to check the ECAL response using the recoil jet in Drell-Yan events (see note in preparation [12]);

5 Remarks on the pilot run at 900 GeV

It is expected that the pilot run at 900 GeV will accumulate a luminosity of 1 to 10 pb^{-1}. The HEEP group will use this opportunity to test the Data Stream (with possibly relaxed cuts). However, these events are not expected to allow for significant physics analyses, as shown in Table 7, which presents the expected numbers of Drell-Yan events with $M_{ee} > 40$ and $100\ \mathrm{GeV/c^2}$.

Drell-Yan events ($\mathrm{GeV/c^2}$)	M >40	M >100
cross section (pb)	42	1.1
nb. events for $10\ \mathrm{pb}^{-1}$	420	11

Table 7: Number of expected Drell-Yan electron pairs with $M_{ee} > 40$ and $100\ \mathrm{GeV/c^2}$ at $\sqrt{s} = 900$ GeV, with a luminosity of $10\ \mathrm{pb}^{-1}$.

The pilot run will however permit a first study of the characteristics of the background and of the fake electrons / photons rates, in order to prepare for the 14 TeV run.

6 Summary and conclusion

The goal of the present note is to prepare the search, at the beginning of CMS data taking, for heavy resonances beyond the Standard Model, decaying into an electron or a photon pair with invariant mass above $1\ \mathrm{TeV/c^2}$.

The proposed strategy relies significantly on the measurement of EW processes ($q\bar{q} \rightarrow e^+e^-$, $q\bar{q} \rightarrow \gamma\gamma$ and $gg \rightarrow \gamma\gamma$) in two energy regions: (i) di-electron pair production at the Z pole, and (ii) di-electron and di-photon pair production with $200 < M < 800$ $\mathrm{GeV/c^2}$; another process to be investigated is Zγ production. These studies are aimed at monitoring the trigger, the particle identification and the detector response, and at estimating systematic uncertainties.

Different possible steps and studies needed for the search of new resonances and for the measurement of the relevant EW processes have been sketched, in particular for the trigger path definition, efficiency measurements and background estimates.

This work will be performed in close coordination with the e$/\gamma$ POG, with the Electroweak and Online Selection Physics Analysis Groups, and with other contributors to the High Energy Pairs study group of the SUSY/BSM PAG.

References

[1] CMS Collaboration, *CMS Physics - Technical Design Report, Volume 2 - Physics Performance*, CERN-LHCC-2006-021, 2006

[2] HEEP group Web page: https://twiki.cern.ch/twiki/bin/view/CMS/SUSYBSMHeep

[3] T. Sjostrand, S. Mrenna and P. Skands, PYTHIA 6.4 physics and manual, Journal of High Energy Physics, Issue 05, pp. 026 (2006)

[4] CMS Collaboration, *CMS Physics - Technical Design Report, Volume 1 - Detector Performance and Software*, CERN-LHCC-2006-001, 2006

[5] Bristol, LLR and RAL, *Effect of ECAL electrons saturation on L1 trigger and reconstruction*, **CMS NOTE** to be submitted

[6] B. Clerbaux, Sh. Elgammal, T. Mahmoud and P. Marage, *Saturation and energy corrections for TeV electrons and photons*, **CMS NOTE** 2006/149

[7] L. Agostino and M. Pieri, *High level trigger selection of electrons and photons*, **CMS NOTE** 2006/078

[8] CMS T0 RTAG, *A T0 Architecture for the CMS Experiment*, **CMS NOTE** 2006/095

[9] Bristol and RAL, note in preparation

[10] Brussels, note in preparation

[11] Les Houches effort reference

[12] Florida, note in preparation