Chapter 1

Data Analysis

The signature of GGM SUSY particle production in this search is an excess of two-photon events with high $\not\!\!E_T$. $\not\!\!E_T$ is reconstructed using the particle flow algorithm as described in Sec. **??**. Candidate two-photon events, as well as control events, are selected according to the offline object criteria presented in Secs. **??** and **??**, the event quality criteria in Sec. **??**, and the trigger requirements in Sec. **??**. These are summarized in Table 1.1.

Table 1.1: Selection criteria for $\gamma\gamma$, $e\gamma$, ee, and $f\!f$ events.

Variable	Cut			
	$\gamma\gamma$	$e\gamma$	ee	$f\!f$
HLT match	IsoVL	IsoVL	IsoVL	IsoVL || R9Id
E_T	$> 40/$ > 25 GeV	$> 40/$ > 25 GeV	$> 40/$ > 25 GeV	$> 40/$ > 25 GeV
SC $\lvert\eta\rvert$	< 1.4442	< 1.4442	< 1.4442	< 1.4442
H/E	< 0.05	< 0.05	< 0.05	< 0.05
$R9$	< 1	< 1	< 1	< 1
Pixel seed	No/No	Yes/No	Yes/Yes	No/No
I_{comb}, $\sigma_{i\eta i\eta}$	< 6 GeV && < 0.011	< 6 GeV && < 0.011	< 6 GeV && < 0.011	< 20 GeV && (≥ 6 GeV || ≥ 0.011)
JSON	Yes	Yes	Yes	Yes
No. good PVs	≥ 1	≥ 1	≥ 1	≥ 1
ΔR_{EM}	> 0.6	> 0.6	> 0.6	> 0.6
$\Delta\phi_{\mathrm{EM}}$	≥ 0.05	≥ 0.05	≥ 0.05	≥ 0.05

This search utilizes 4.7 fb^{-1} of CMS data collected during the period April-December 2011, corresponding to the following datasets [?]:

- `/Photon/Run2011A-05Jul2011ReReco-ECAL-v1/AOD`
- `/Photon/Run2011A-05Aug2011-v1/AOD`
- `/Photon/Run2011A-03Oct2011-v1/AOD`
- `/Photon/Run2011B-PromptReco-v1/AOD`

The search strategy is to model the backgrounds to the GGM SUSY signal using $\not\!\!E_T$ shape templates derived from the control samples, and then to look for a high-$\not\!\!E_T$ excess above the estimated background in the $\gamma\gamma$ sample. There are two categories of backgrounds: QCD processes with no real $\not\!\!E_T$ and electroweak processes with real $\not\!\!E_T$ from neutrinos. The relevant QCD background processes are multijet production with at least two jets faking photons, photon + jet production with at least one jet faking a photon, diphoton production, and Z production with a radiated photon where at least one of the Z decay products (typically a jet) fakes a photon. The relevant electroweak background processes, which are small compared to the QCD background, involve $W \to e\nu$ decay with a recoiling jet that fakes a photon or a real radiated photon. In both cases, the electron is misidentified as a photon due to a small inefficiency in reconstructing the electron pixel seed. The main diagrams contributing to the QCD(electroweak) backgrounds are shown in Figure **??**(**??**). **Generate these Feynman diagrams.**

Figure **??** shows the $\not\!\!E_T$ spectrum of the $\gamma\gamma$ search data sample overlaid on the $\not\!\!E_T$ spectra of MC simulated background components. The MC spectra are normalized to the integrated luminosity of the $\gamma\gamma$ data sample. **Make this plot.** The dominant background components are QCD inclusive photon processes. The MC is not used in the actual background estimation. It is just shown here to illustrate the breakdown of backgrounds.

Data control samples are used to model all of the backgrounds. The primary control sample used to model the QCD background is the *ff* sample, which is similar to the candidate $\gamma\gamma$ sample but with combined isolation or $\sigma_{i\eta i\eta}$ cuts inverted. The cuts on these variables are used to distinguish between photons and jets, so by inverting those cuts, the resulting *ff* sample becomes enriched with QCD dijets. Because the fake photons are still required to pass a tight cut on H/E, they are guaranteed to be very electromagnetic jets, with an EM energy scale and resolution similar to that of the candidate photons. This insures that the resulting estimate of the $\not\!\!E_T$ shape does not have too long of a tail from severe HCAL mis-measurements that are actually rare in the $\gamma\gamma$ sample, as shown in Figure **??**. **Plot the $\gamma\gamma/ff$ $\not\!\!E_T$ agreement for different values of the ff H/E cut in MC. Make the same plot in data for a restricted $\not\!\!E_T$ range?**

As a cross-check, the *ee* sample is also used to model the QCD background. This sample of Z decays should have no true $\not\!\!E_T$, just like the *ff* sample, and the electron definition (differing from the photon definition only in the presence of a pixel seed) insures that the electron energy scale and resolution is similar to that of the photon.

Finally, the $e\gamma$ sample is used to model the electroweak background from $W \to e\nu$ decays. The $e\gamma$ $\not\!\!E_T$ distribution is scaled by the electron→photon misidentification rate to predict the number of $W\gamma$ and W + jet events in the $\gamma\gamma$ sample.

The remainder of this chapter describes the data analysis procedures and the final results of the search. Sec. 1.1 addresses the QCD background estimation. Sec. **??** addresses the electroweak background estimation. The chapter concludes with a discussion of systematic errors in Sec. 1.3 and a presentation of the final results in Sec. 1.4.

1.1 Modeling the QCD Background

1.1.1 Outline of the Procedure

Due to the fact that the CMS ECAL energy resolution is much better than the HCAL energy resolution, the energies of the two candidate photons in the events of the $\gamma\gamma$ sample are typically measured to far greater accuracy and precision than the energy of the hadronic recoil in those events. Therefore, fake $\not\!\!E_T$ in the $\gamma\gamma$ sample is almost entirely the result of hadronic mis-measurement in QCD dijet, photon + jet, and diphoton events. The strategy employed to model this background is to find a control sample in data consisting of two well-measured EM objects, just like the candidate $\gamma\gamma$ sample, and assign each event a weight to account for the underlying kinematic differences between the control and candidate samples. Once the reweighted $\not\!\!E_T$ spectrum of the control sample is created, it is then normalized in the low-$\not\!\!E_T$ region, the assumption being that GGM SUSY does not predict a significant amount of events at low $\not\!\!E_T$. There are three aspects to this QCD background estimation procedure that bear highlighting:

Choice of control sample Since the underlying cause of $\not\!\!E_T$ in the candidate sample is mis-measured hadronic activity, a control sample with similar hadronic activity to the candidate sample should be chosen. Hadronic activity refers to number of jets, jet E_T, pileup, etc.

Reweighting The control sample is reweighted so that its $\not\!\!E_T$ spectrum appears as it would if the control sample had the same kinematic properties as the candidate sample (i.e. particle p_T and η distributions, etc.). By choosing an appropriate control sample and reweighting it, the control $\not\!\!E_T$ distribution should now match both the hadronic activity and the kinematics of the candidate sample.

Normalization Finally, the control $\not\!\!E_T$ distribution is normalized in a region of low $\not\!\!E_T$, where contamination from the expected GGM SUSY signal is small. This implies an extrapolation of the low-$\not\!\!E_T$ QCD background prediction to the high-$\not\!\!E_T$ signal region.

As explained in the beginning of this chapter, the *ff* sample is used as the primary QCD control sample, while the *ee* sample is used as a cross-check. Both samples have two well-measured EM objects per event, no real $\not\!\!E_T$, and similar hadronic activity to the $\gamma\gamma$ sample. Figure **??** shows a comparison of the shapes of some distributions relevant to hadronic activity between the $\gamma\gamma$, *ee*, and *ff* samples. **Make an observation about the lesser hadronic activity in the *ee* sample and how the reweighting procedure will account for that.**

Table 1.2: Definition of hadronic jets. **Add a footnote describing the PF electron and PF muon definitions, with references.**

Variable	Cut
Algorithm	`L1FastL2L3Residual` corrected PF (cf. Sec. **??**)
p_T	> 30 GeV
$\lvert\eta\rvert$	< 5.0
Neutral hadronic energy fraction	< 0.99
Neutral electromagnetic energy fraction	< 0.99
Number of constituents	> 1
Charged hadronic energy	> 0.0 GeV if $\lvert\eta\rvert < 2.4$
Number of charged hadrons	> 0 if $\lvert\eta\rvert < 2.4$
Charged electromagnetic energy fraction	< 0.99 if $\lvert\eta\rvert < 2.4$
ΔR to nearest electron, muon, or one of the two primary EM objects	> 0.5

1.1.2 Reweighting

To reweight the control sample events to match the kinematics of the candidate sample events, a weight based on the p_T of the di-EM-object system and the number of jets

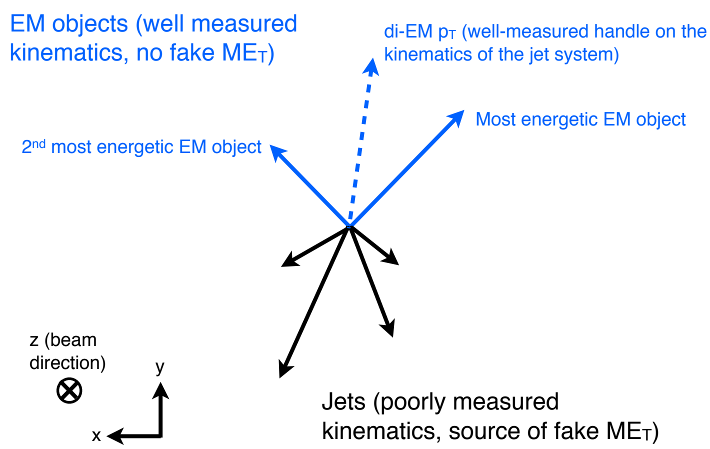


Figure 1.1: Cartoon showing the di-EM system in blue and the hadronic recoil in black. The di-EM p_T (dashed blue line) is used to reweight the control sample kinematic properties to match those of the candidate $\gamma\gamma$ sample.

.

in the event is used. As explained in Sec. 1.1.1, $\not\!\!E_T$ in the $\gamma\gamma$, $f\!f$, and ee samples is due to the poorly measured hadronic recoil off the well-measured di-EM system. Therefore, the p_T of the di-EM system is a good handle on the true magnitude of the hadronic recoil, which affects the measured $\not\!\!E_T$. The di-EM system is depicted in Figure 1.1.

Whereas the di-EM p_T reweighting accounts for differences in production kinematics between the control and $\gamma\gamma$ samples, a simultaneous reweighting based on the number of jets in the event accounts for differences in hadronic activity between the samples, especially between ee and $\gamma\gamma$ (cf. Fig. **??**). Jets are defined as in Table 1.2, but with $|\eta|$ restricted to 2.6 (i.e. HF jets excluded). Figure **??** shows the effect of reweighting by number of jets per event, which is to increase(decrease) the tail of the ee($f\!f$) $\not\!\!E_T$ spectrum.

Although the electron and photon energies are well measured by the ECAL, the ECAL-only measurement of the fake photon energy (cf. Sec **??**) is biased slightly low due to the fact that fakes (which are really jets) tend to deposit some energy in the HCAL. This can be seen in Figs. 1.2 and 1.3, which show the relative difference between the ECAL-only E_T measurement and the PF E_T measurement vs. EMF for electrons, photons, and fakes. PF E_T is defined as the `L1Fast`-corrected E_T of the nearest PF jet with $p_T \geq 20$ GeV (i.e., the E_T of the PF jet object reconstructed from the same ECAL shower as the fake photon). On average, the fakes tend to deposit a few percent more energy in the HCAL than the electrons or photons, which is recovered by the PF algorithm. For this reason, the PF p_T is used in the calculation of di-EM p_T rather than the ECAL-only p_T. This leads to a modest improvement in the agreement between the *ee* and *ff* $\not{E}_T$ spectra, as shown in Figure **??**.

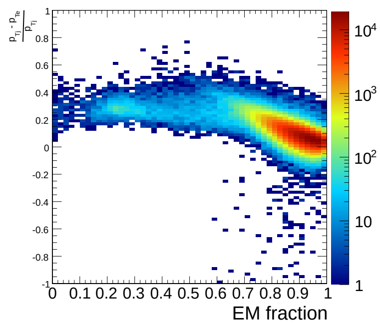
The control sample event weights are defined as

$$w_{ij} = \frac{N_{\text{control}}}{N_{\gamma\gamma}}\frac{N^{ij}_{\gamma\gamma}}{N^{ij}_{\text{control}}} \tag{1.1}$$

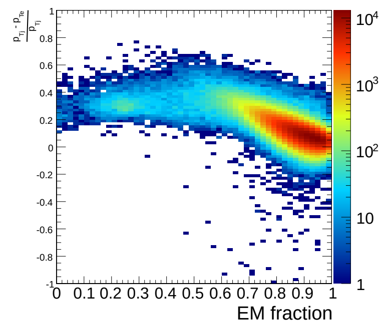
where i runs over the number of di-EM p_T bins, j runs over the number of jet bins, N_{control} is the total number of events in the control sample, $N_{\gamma\gamma}$ is the total number of events in the $\gamma\gamma$ sample, $N^{ij}_{\gamma\gamma}$ is the number of $\gamma\gamma$ events in the i^{th} di-EM p_T bin and j^{th} jet bin, and N^{ij}_{control} is the number of control sample events in the i^{th} di-EM p_T bin and j^{th} jet bin. The effect of the reweighting is more significant for the *ee* sample than for the *ff* sample, as shown in Figure **??**.

1.1.3 Normalization

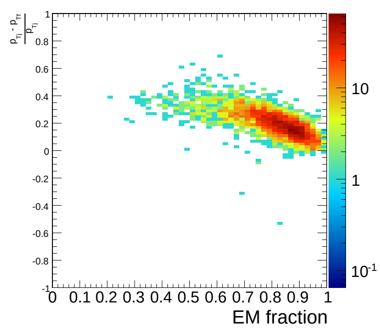
After reweighting, the $\not{E}_T$ distributions of the QCD control samples are normalized to the $\not{E}_T < 20$ GeV region of the candidate $\gamma\gamma$ $\not{E}_T$ spectrum, where signal contamination is low. The normalization factor is $(N^{\not{E}_T<20\text{GeV}}_{\gamma\gamma} - N^{\not{E}_T<20\text{GeV}}_{\text{electroweak}})/N^{\not{E}_T<20\text{GeV}}_{\text{control}}$, where $N^{\not{E}_T<20\text{GeV}}_{\text{electroweak}}$ is the expected number of electroweak background events with $\not{E}_T < 20$ GeV (discussed in Section 1.2).



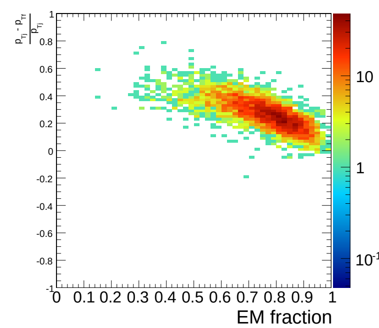
(a) Leading electron in ee events.



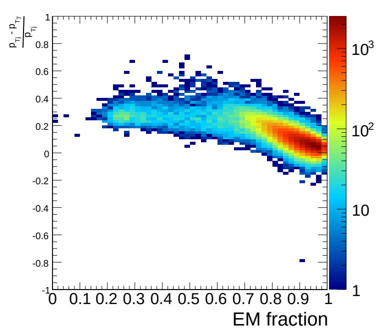
(b) Trailing electron in ee events.



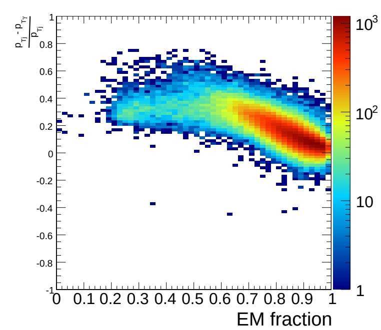
(c) Leading fake in ff events.



(d) Trailing fake in ff events.



(e) Leading photon in $\gamma\gamma$ events.



(f) Trailing photon in $\gamma\gamma$ events.

Figure 1.2: Relative difference between the ECAL-only E_T measurement and the PF E_T measurement vs. EMF. PF E_T is defined in the text. **Replace with current figure.**

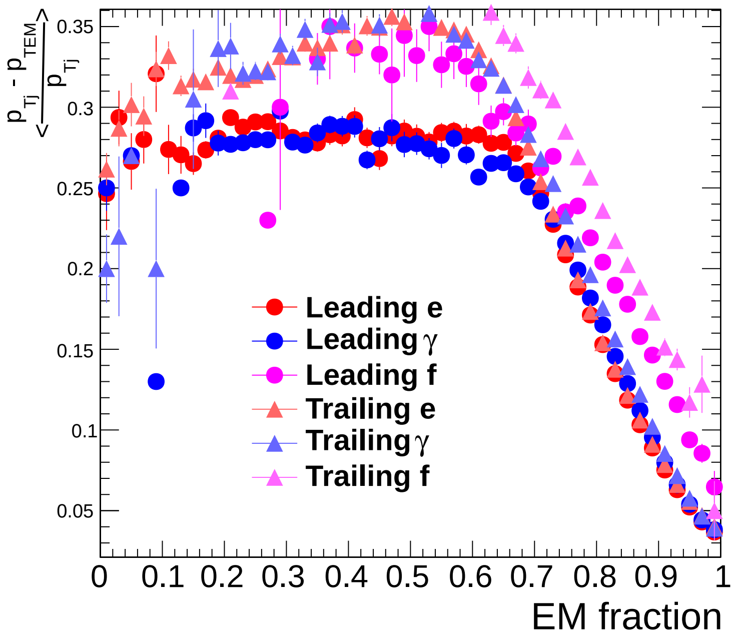


Figure 1.3: Average relative difference between the ECAL-only E_T measurement and the PF E_T measurement vs. EMF for the two electrons in ee events, the two fakes in ff events, and the two photons in $\gamma\gamma$ events (i.e. profile histograms of Fig. 1.2). PF E_T is defined in the text. **Replace with current figure.**

1.1.4 ff Control Sample

1.2 Modeling the Electroweak Background

1.3 Systematic Errors

1.3.1 Jet Energy Scale Uncertainty

The dijet $\mathrm{p_T}$ reweighting method utilizes jets corrected for imperfect calorimeter response (see Sec. **??** for a description of the jet reconstruction and correction procedure). Since the applied jet energy scale (JES) factor has an error associated to it due to the limitations of the JES derivation ([**?**] and Sec. **??**), this uncertainty must be propagated to the uncertainty on the dijet $\mathrm{p_T}$ weights.

The JES contribution to the dijet $\mathrm{p_T}$ weights is estimated by performing 1000 pseudo-experiments on each of the $\gamma\gamma$ and ff samples. For the purpose of estimating the JES error, the results of the true experiment may be thought of as a set of measurements:

- The set of **uncorrected jet 4-vectors** corresponding to the **leading EM object** in the $\gamma\gamma$ sample $\left\{\mathrm{p}_{\mathrm{j1}}^{\mu 1}, \mathrm{p}_{\mathrm{j1}}^{\mu 2}, ..., \mathrm{p}_{\mathrm{j1}}^{\mu \mathrm{N}_{\gamma\gamma}}\right\}$
- The set of **uncorrected jet 4-vectors** corresponding to the **trailing EM object** in the $\gamma\gamma$ sample $\left\{\mathrm{p}_{\mathrm{j2}}^{\mu 1}, \mathrm{p}_{\mathrm{j2}}^{\mu 2}, ..., \mathrm{p}_{\mathrm{j2}}^{\mu \mathrm{N}_{\gamma\gamma}}\right\}$
- The set of **JES** accompanying the uncorrected jet 4-vectors corresponding to the **leading EM object** in the $\gamma\gamma$ sample $\left\{\mathrm{c}_{\mathrm{j1}}^{1}, \mathrm{c}_{\mathrm{j1}}^{2}, ..., \mathrm{c}_{\mathrm{j1}}^{\mathrm{N}_{\gamma\gamma}}\right\}$
- The set of **JES** accompanying the uncorrected jet 4-vectors corresponding to the **trailing EM object** in the $\gamma\gamma$ sample $\left\{\mathrm{c}_{\mathrm{j2}}^{1}, \mathrm{c}_{\mathrm{j2}}^{2}, ..., \mathrm{c}_{\mathrm{j2}}^{\mathrm{N}_{\gamma\gamma}}\right\}$
- The set of **JES uncertainties** accompanying the uncorrected jet 4-vectors corresponding to the **leading EM object** in the $\gamma\gamma$ sample $\left\{\sigma_{\mathrm{cj1}}^{1}, \sigma_{\mathrm{cj1}}^{2}, ..., \sigma_{\mathrm{cj1}}^{\mathrm{N}_{\gamma\gamma}}\right\}$
- The set of **JES uncertainties** accompanying the uncorrected jet 4-vectors corresponding to the **trailing EM object** in the $\gamma\gamma$ sample $\left\{\sigma_{\mathrm{cj2}}^{1}, \sigma_{\mathrm{cj2}}^{2}, ..., \sigma_{\mathrm{cj2}}^{\mathrm{N}_{\gamma\gamma}}\right\}$
- The set of **uncorrected jet 4-vectors** corresponding to the **leading EM object** in the ff sample $\left\{\mathrm{p}_{\mathrm{j1}}^{\mu 1}, \mathrm{p}_{\mathrm{j1}}^{\mu 2}, ..., \mathrm{p}_{\mathrm{j1}}^{\mu \mathrm{N}_{\mathrm{ff}}}\right\}$
- The set of **uncorrected jet 4-vectors** corresponding to the **trailing EM object** in the ff sample $\left\{\mathrm{p}_{\mathrm{j2}}^{\mu 1}, \mathrm{p}_{\mathrm{j2}}^{\mu 2}, ..., \mathrm{p}_{\mathrm{j2}}^{\mu \mathrm{N}_{\mathrm{ff}}}\right\}$
- The set of **JES** accompanying the uncorrected jet 4-vectors corresponding to the **leading EM object** in the ff sample $\left\{\mathrm{c}_{\mathrm{j1}}^{1}, \mathrm{c}_{\mathrm{j1}}^{2}, ..., \mathrm{c}_{\mathrm{j1}}^{\mathrm{N}_{\mathrm{ff}}}\right\}$
- The set of **JES** accompanying the uncorrected jet 4-vectors corresponding to the **trailing EM object** in the ff sample $\left\{\mathrm{c}_{\mathrm{j2}}^{1}, \mathrm{c}_{\mathrm{j2}}^{2}, ..., \mathrm{c}_{\mathrm{j2}}^{\mathrm{N}_{\mathrm{ff}}}\right\}$

- The set of **JES uncertainties** accompanying the uncorrected jet 4-vectors corresponding to the **leading EM object** in the ff sample $\left\{\sigma^{1}_{\mathrm{cj1}}, \sigma^{2}_{\mathrm{cj1}}, ..., \sigma^{\mathrm{N_{ff}}}_{\mathrm{cj1}}\right\}$
- The set of **JES uncertainties** accompanying the uncorrected jet 4-vectors corresponding to the **trailing EM object** in the ff sample $\left\{\sigma^{1}_{\mathrm{cj2}}, \sigma^{2}_{\mathrm{cj2}}, ..., \sigma^{\mathrm{N_{ff}}}_{\mathrm{cj2}}\right\}$

From these measurements, the $\gamma\gamma$ and ff dijet $\mathrm{p_T}$ spectra and the resulting ff dijet weights can be calculated. In each of the 1000 pseudo-experiments, a new set of JES factors is generated according to the measured JES uncertainties, and new dijet $\mathrm{p_T}$ spectra and weights are subsequently calculated. The spread of the 1000 weights (binned in dijet $\mathrm{p_T}$) is taken as the error due to JES uncertainty. The total error on the weights is the quadrature sum of the JES error and the statistical error, and is propagated to the error on the final $\not\!\!E_T$ measurement via a similar pseudo-experiment procedure described in Sec. 1.3.2.[1]

If the JES uncertainty were to cause the jet energy to be reconstructed below the 20 GeV ntuple cut, there could be a small error or bias in the $\not\!\!E_T$ introduced due to EM-matched jets falling below the matching threshold. The percentage of jets lost due to jet $\mathrm{E_T}$ matching threshold has been checked in data, and found to be X% (X% of events). Furthermore, the trailing EM $\mathrm{E_T}$ cut is 25 GeV/c, implying that the JES would have to be mis-measured by at least 20% to fall below the jet matching threshold. Since the typical JES uncertainty is no more than 5%, a mis-measurement of this type is a 4σ event and should occur in only 0.1% of cases. As expected, this effect is negligible, as shown in Figure X.

1.3.2 Statistical Uncertainty in the ff or ee Weights

1.4 Results

[1]The $\not\!\!E_T$ is uncorrected and therefore its central value per event is unaffected by a change in the JES.