Electromagnetic showers reconstruction and analysis and neutrino oscillation study by electron detection

Amina Zghiche and Florian Brunet

Laboratoire d'Annecy-le-vieux de Physique des Particules - LAPP

Réunion de groupe
November 26th, 2012

Contents

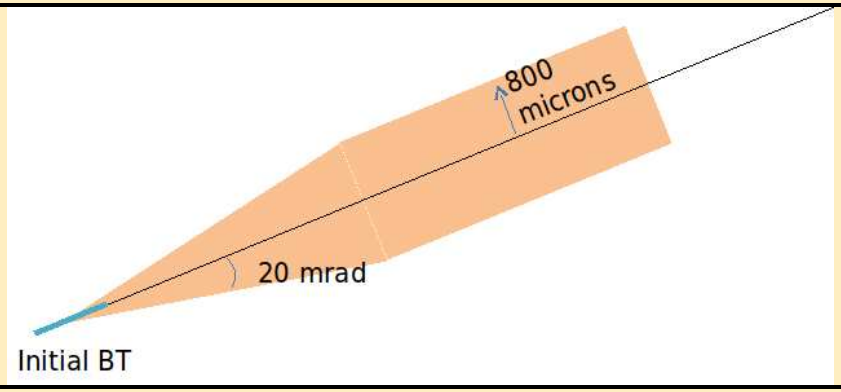
Contents

Electromagnetic Shower Reconstruction Algorithm

Principle of the shower reconstruction

1. The shower initiation with reconstructed tracks
2. The shower reconstruction : connecting basetracks from a film to the next one based on geometrical criteria.

Summary of selections in the basetrack collection process

Values of the selections
$abs(x_{btk} - x_{init}) \leq 1250.$ AND $abs(y_{btk} - y_{init}) \leq 1250.$
$abs(\theta_X) < 0.5\,rd$ AND $abs(\theta_Y) < 0.5\,rd$
$abs(\theta^{3D}_{btk} - \theta^{3D}_{init}) < 0.150\,rd$
$abs(r_{btk} - r_{init}) < 0.150\,\mu m$


Electromagnetic Shower Reconstruction Algorithm

Outline of Identification and Energy Estimation

Once showers are reconstructed, some parameters relevant for electron identification and energy estimation are extracted and used to feed an Artificial Neural Network (ANN) for both purposes.

Description of the variables	Typical value for electrons at 4 GeV	Typical value for pions at 4 GeV
Multiplicity of basetracks in the shower	21.3 ± 1.8	12.8 ± 0.2
The mean value of the angular difference between an initiator and one basetrack in the shower distribution	35.9 ± 0.2 mrad	3.97 ± 0.01 mrad
The RMS value of the angular difference between an initiator and one basetrack in the shower distribution	24.9 ± 0.1 mrad	2.87 ± 0.01 mrad
The mean value of the position difference between an initiator and one basetrack in the shower distribution	$40.1 \pm 0.2\ \mu m$	$10.64 \pm 0.03\ \mu m$
The RMS value of the position difference between an initiator and one basetrack in the shower distribution	$27.1 \pm 0.1\ \mu m$	$7.93 \pm 0.03\ \mu m$
The content of each bin of the shower longitudinal profile i.e. the number of basetracks per film.	-	-

Electromagnetic Shower Reconstruction Algorithm

Reconstruction process

Each MC event is processed through the reconstruction chain :

- Muon identification : OpCarac
- Brick finding
- CS simulation
- Scanback procedure
- Totalscan : MT linking, films alignment, BT linking, Vertexing
- Shower reconstruction applied to all reconstructed tracks in the Totalscan volume

Shower selection

Each event will contain several reconstructed showers, one has to select the most probable electron shower

- Select only showers identified by the ANN as electron. (eProb1==1)
- Among those selected above, keep the most upstream showers.
- Among those selected above, the shower with the largest number of BT.

Electromagnetic Shower Reconstruction Algorithm

Recent upgrades of the shower algorithm

Shower algorithm implementation in OpEmuRec has been debugged and comes to modifications committed to fedra SVN as well as OpEmuRec CVS :

- Thickness of plates in the OpEmuRec MC was different from the one used in the shower algo
- Bug in the number of plates used to extract shower profiles for events spanned over 45 plates and more
- Static arrays not sufficient to store large BT multiplicity events (ν_e prompt)
- Propagation of the MC true information related to basetrack to the output of the shower algorithm

A documentation of the shower algorithm usage with OpEmuRec is available on the CVS server.

MC samples used in this study

Description of all MC samples.

Name	Channel	Statistics
MC_nue_1k	ν_e from the beam contamination	1000
MC_oscillated_nue_1k	oscillated ν_e	1000
MC_taue_3k	oscillated ν_τ with τ decay in electronic mode	3000
MC_numunc_3k	ν_μ NC interactions	3000
MC_numucc_3k	ν_μ CC interactions	3000
MC_electronPion_4GeV_1k	Mixture of electrons and pions	1000
MC_electronPion_2GeV_1k	Mixture of electrons and pions	1000
MC_electron_10k	Electrons	10000

Contents

Electron identification

The software framework is OpRelease 4.0.

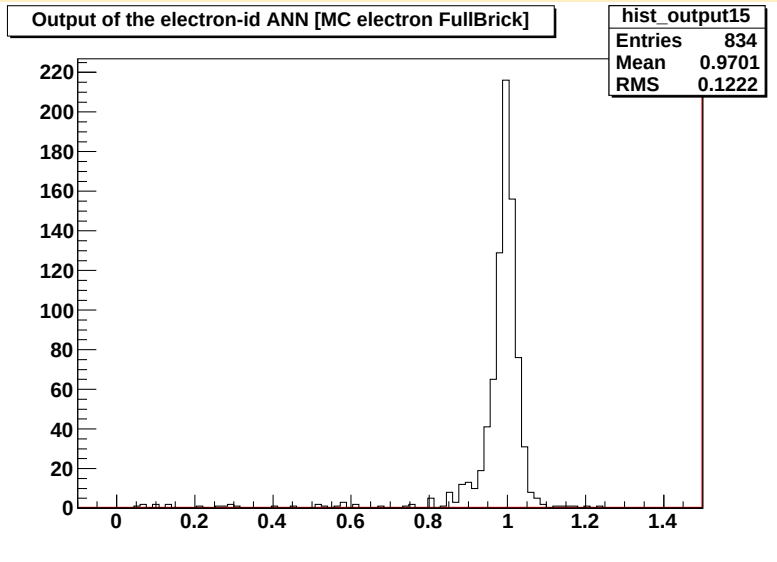
Strategy to assess performance and efficiency of the identification.

- Define a cut on the ANN output to identify electrons.
- Assess identification performance with testbeam data and compare it with a dedicated MC sample "testbeam-like".
- Evaluate identification efficiency with MC simulations and estimate its systematic uncertainty with 1000 ν_e events coming from the oscillation and 3000 ν_τ events coming from the oscillation with the tau in its electronic decay.
- Check the identification efficiency dependancies with the energy and the vertex depth.

Performance of electron identification

The ANN gives an output variable close to 1 for eletrons and to 0 for pions. An appropriate cut on this value to ensure that less than 1% of the selected showers are from pion contamination.

MC simulation NN output distribution for the pure electron sample using the full brick information.



These distributions allow to define a cut on the ANN output at a value 0.78 beyond which the shower is identified as an electron.

Performance of electron identification

OPERA bricks have been exposed in the summer 2011 at CERN to 2 and 4 GeV pion enriched in electrons beam. These bricks have been then scanned in Bern thanks to Tomoko, Ariga, Serhan and Annika and analysed in Bern thanks to Ariga.

Result of the electron testbeam data processing

E (GeV)	"Beam" tracks	Expected electrons	Selected tracks	Electron tracks confirmed
2	350	80	69	49
4	960	50	52	43

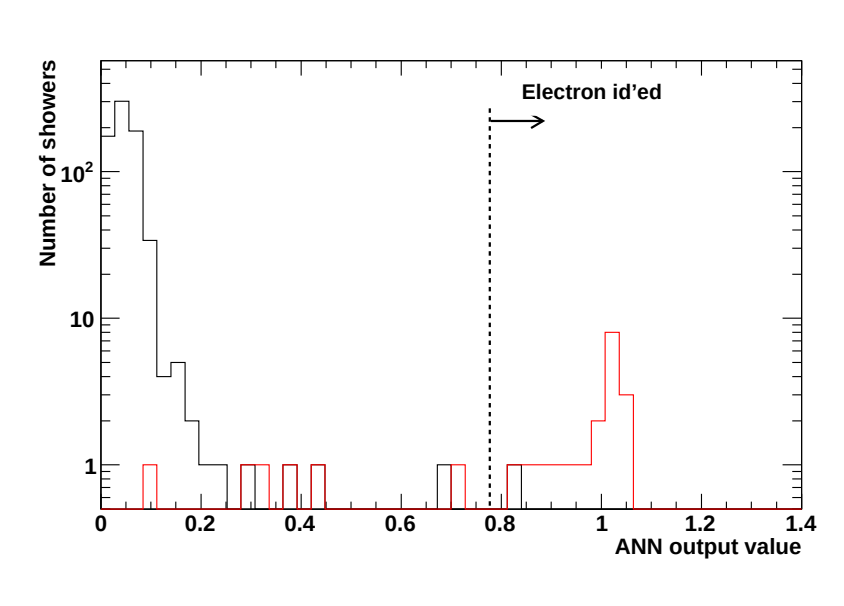
Performance of the algorithm by comparing testbeam data and MC testbeam-like

E (GeV)	MC		DATA		
	$\epsilon_{e\to e}$ (%)	$\epsilon_{\pi\to e}$ (%)	η_e (%)	Purity (%)	κ_e(%)
2	60.2$\pm$4.5	0.0	86 ± 4	71 ± 5	61 ± 5
4	76$\pm$8	0.14	100	83 ± 5	86 ± 5

Table: Summary of the electron identification algorithm performance with $\epsilon_{\alpha\to\beta} = \frac{\text{Number of MC true } \beta \text{ identified as } \alpha}{\text{Total number of MC true } \beta}$ and $\eta_\alpha = \frac{\text{Number of particle identified as } \alpha}{\text{Number of expected } \alpha \text{ particles}}$. An estimation of the purity of the "confirmed" electrons in data and of the $\kappa_e = \frac{\text{\# of confirmed}}{\text{\# of expected}}$ has been checked by visual inspection.

Performance of electron identification

Performance of the algorithm by comparing testbeam data and MC testbeam-like



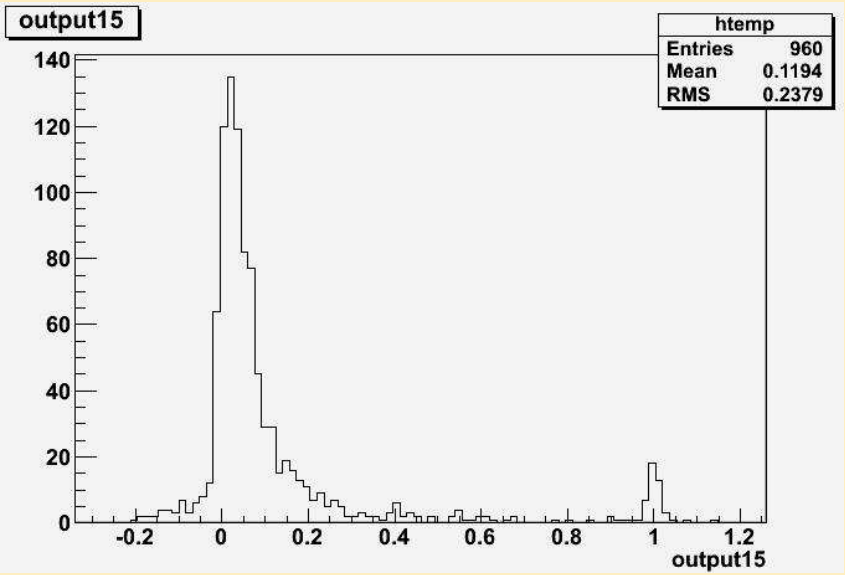
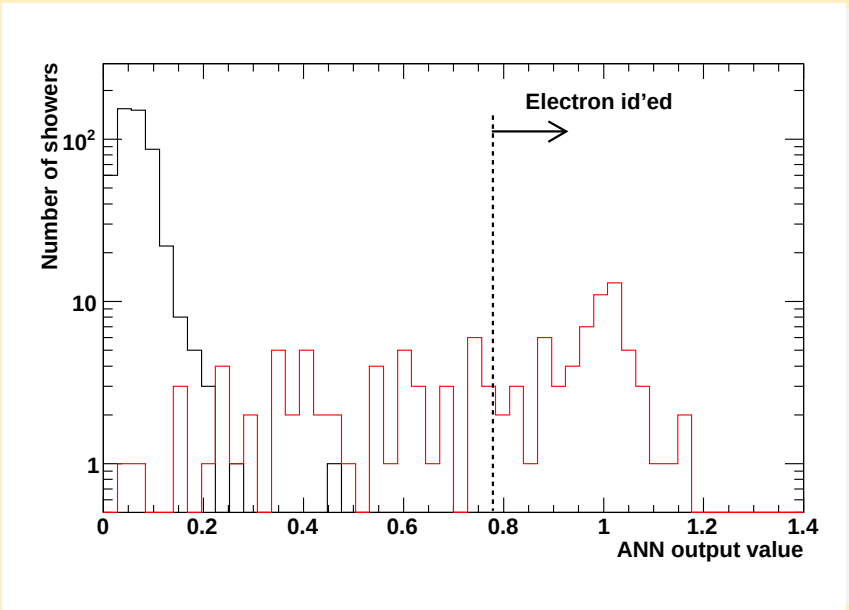


Figure: Output variable of the identification ANN for testbeam data at 4 GeV on the right side [Ariga] and for a MC sample testbeam-like on the left side showing ANN output value for MC true electron in red and for MC true pion in black.

Performance of electron identification

Performance of the algorithm by comparing testbeam data and MC testbeam-like



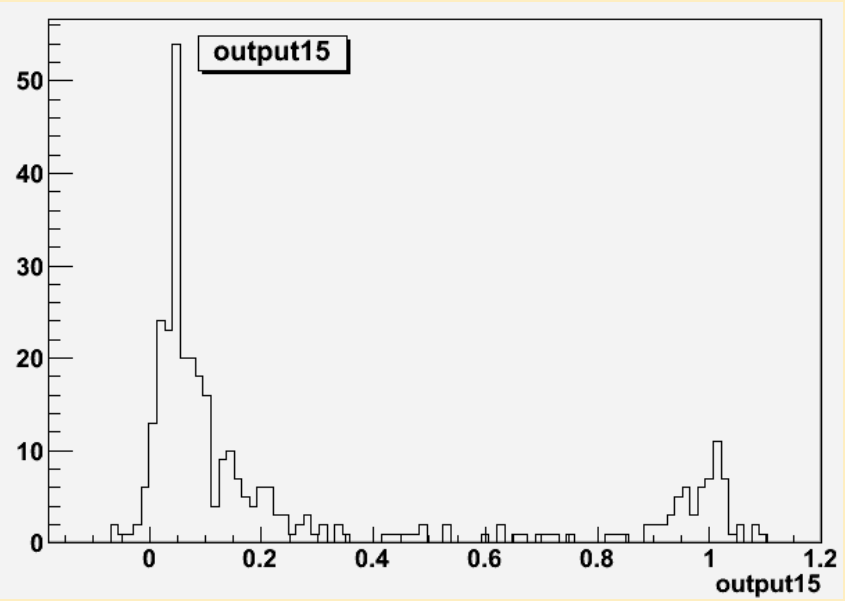


Figure: Output variable of the identification ANN for testbeam data at 2 GeV on the right side [Ariga] and for a MC sample testbeam-like on the left side showing ANN output value for MC true electron in red and for MC true pion in black.

Identification efficiency

Shower reconstruction and electron identification processes thus established are applied to MC simulation of OPERA electron neutrino events, both from $\tau_{\rightarrow}e$ and from the oscillation process $\nu_\mu \rightarrow \nu_e$. The identification efficiency has been evaluated as a function of a variable number of films.

Identification efficiency of the $\nu_\mu \rightarrow \nu_e$ oscillation sample used in the MC simulation

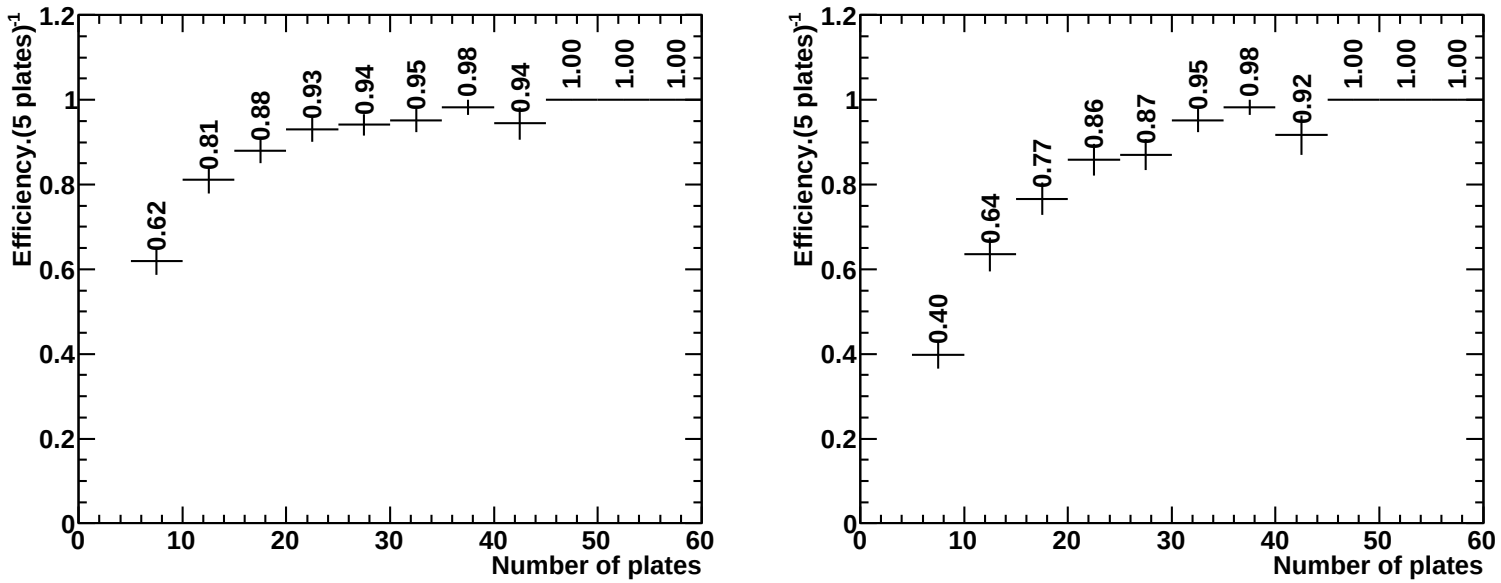
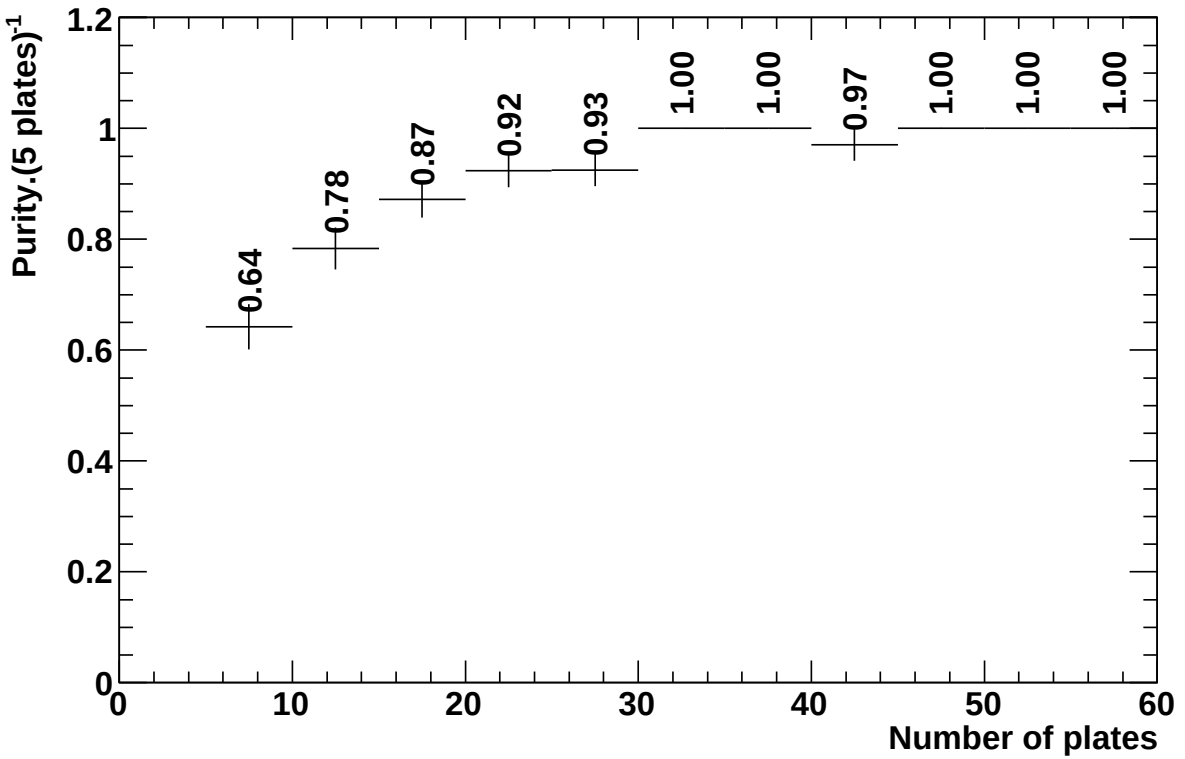


Figure: Electron Identification efficiency as a function of the shower extension in terms of number of films obtained with MC simulation of the $\nu_\mu \rightarrow \nu_e$ oscillation channel on the left panel. The same efficiency but only for MC true electron showers on the right panel. The difference gives the systematic uncertainty on identification of showers.

Identification efficiency

Identification purity of the $\nu_\mu \rightarrow \nu_e$ oscillation sample used in the MC simulation



Identification efficiency

Identification of electron showers with the MC true information in the $\nu_\mu \rightarrow \nu_e$ oscillation channel.

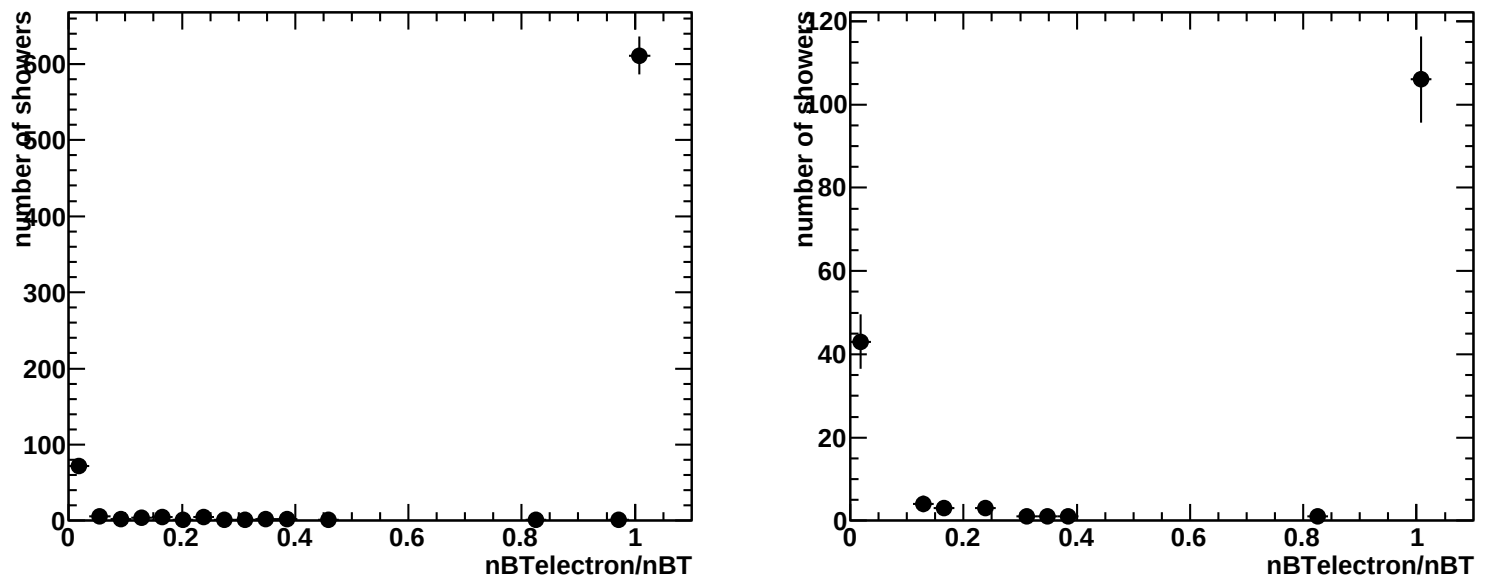


Figure: Ratio of BT linked to a hit of the primary electron divided by the total number of BT in the shower with all events in the left panel and with only showers which have an extension of 10 plates at maximum in the right panel. Both of these plots are obtained with MC simulation of the $\nu_\mu \rightarrow \nu_e$ oscillation channel. A cut to 0.5 is used to evaluate the misidentification of electron showers.

Identification efficiency

Identification efficiency of the $\nu_\mu \rightarrow \nu_\tau(\tau \rightarrow e)$ sample used in the MC simulation

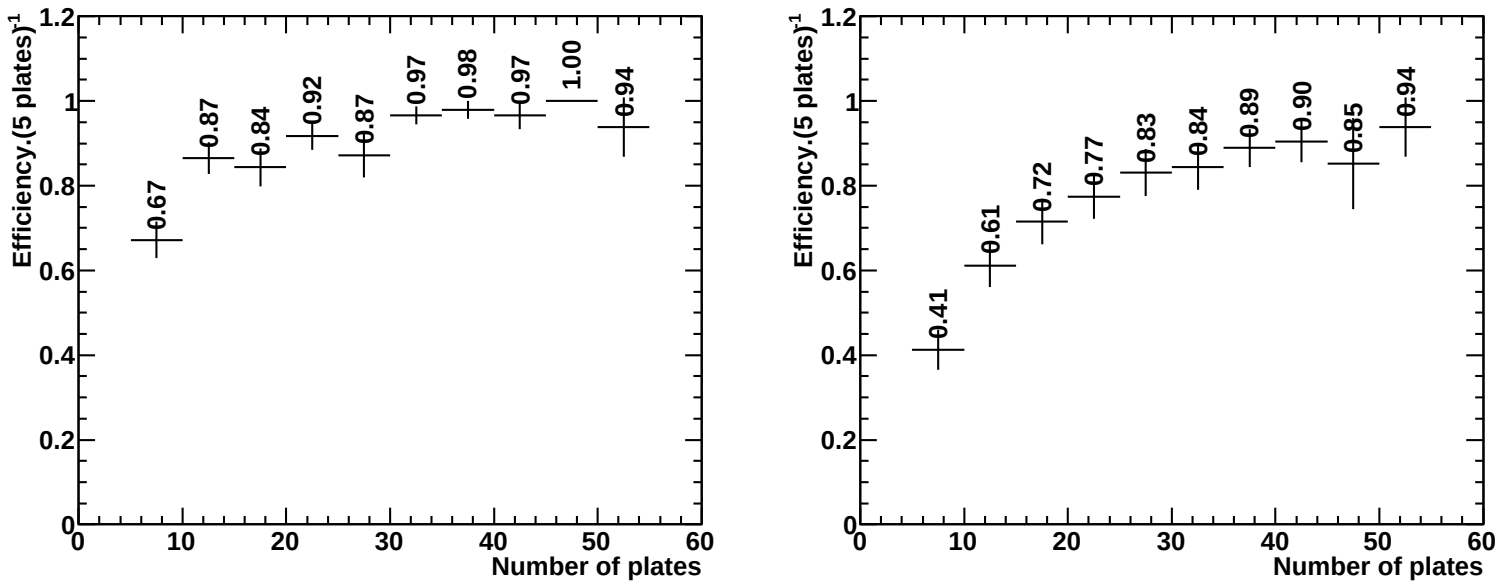
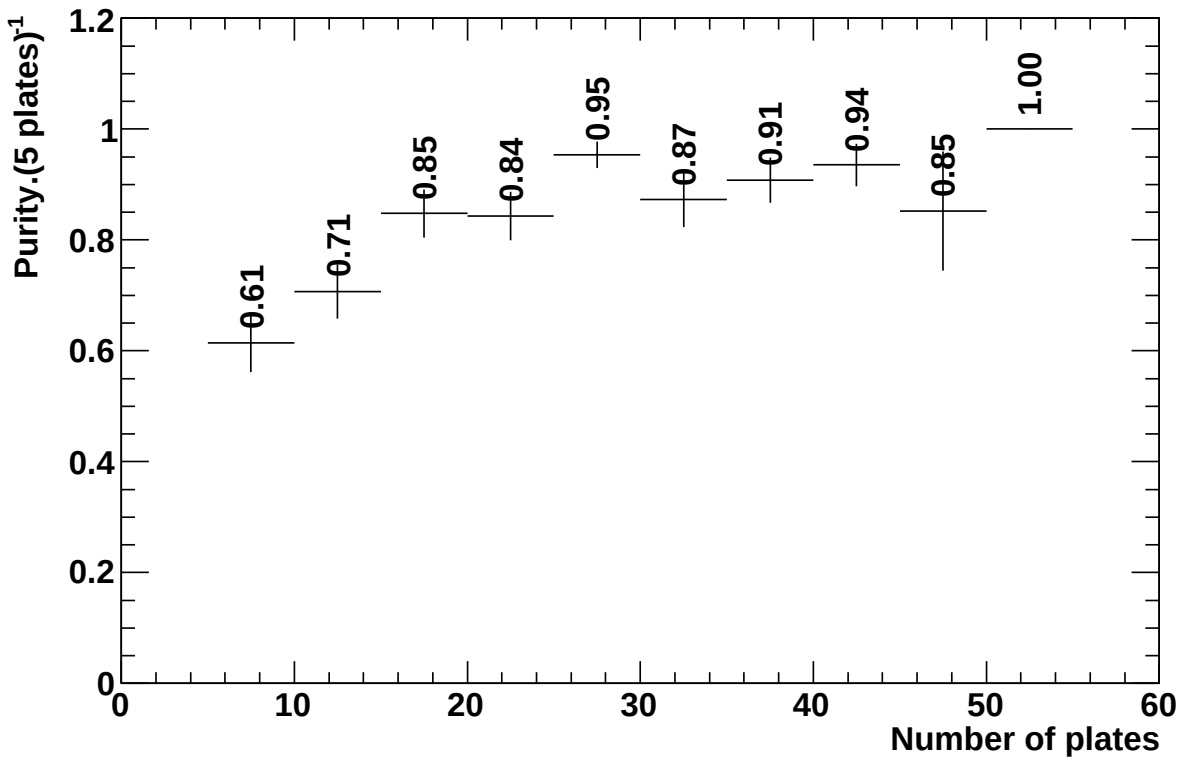


Figure: Electron Identification efficiency as a function of the shower extension in terms of number of films obtained with MC simulation of the $\nu_\mu \rightarrow \nu_e$ oscillation channel on the left panel. The same efficiency but only for MC true electron showers on the right panel. The difference gives the systematic uncertainty on identification of showers.

Identification efficiency

Identification purity of the $\nu_\mu \rightarrow \nu_\tau(\tau \rightarrow e)$ sample used in the MC simulation



Identification efficiency

Identification of electron showers with the MC true information in the $\nu_\mu \rightarrow \nu_\tau(\tau \rightarrow e)$ channel.

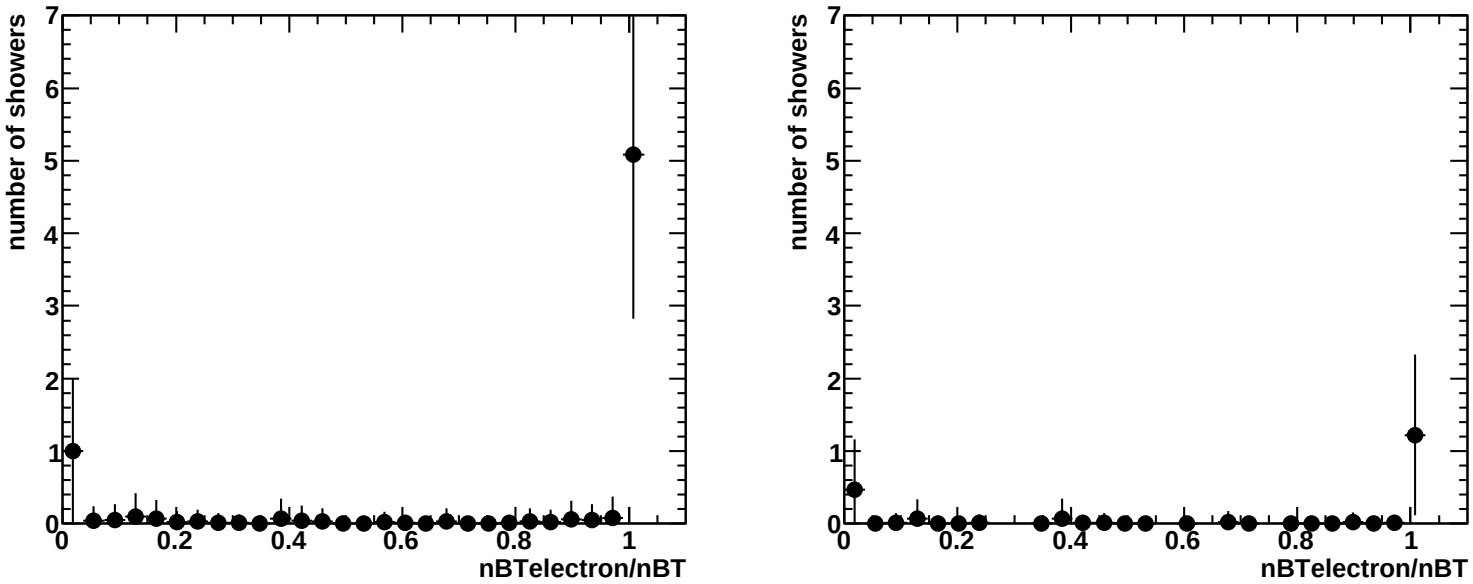


Figure: Ratio of BT linked to a hit of the primary electron divided by the total number of BT in the shower with all events in the left panel and with only showers which have an extension of 10 plates at maximum in the right panel. Both of these plots are obtained with MC simulation of the $\nu_\mu \rightarrow \nu_\tau(\tau \rightarrow e)$ channel. A cut to 0.5 is used to evaluate the misidentification of electron showers.

Identification efficiency dependancies

Identification efficiency dependancy with the energy in the $\nu_\mu \rightarrow \nu_e$ oscillation channel.

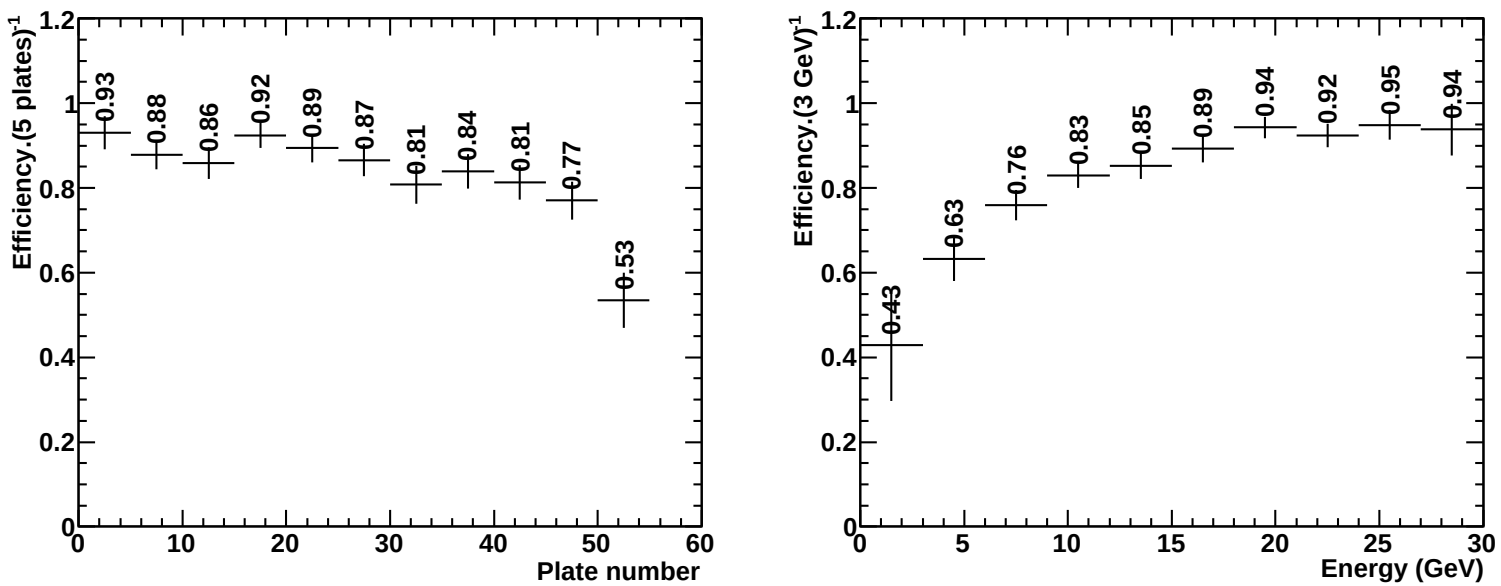


Figure: Electron identification efficiency as a function of the electron MC true energy in the right panel and electron identification efficiency as a function of the MC true vertex depth in the left panel, both obtained with MC simulation of the $\nu_\mu \rightarrow \nu_e$ oscillation channel.

Identification efficiency dependancies

Identification efficiency dependancy with the energy in the $\nu_\mu \rightarrow \nu_\tau(\tau \rightarrow e)$ channel.

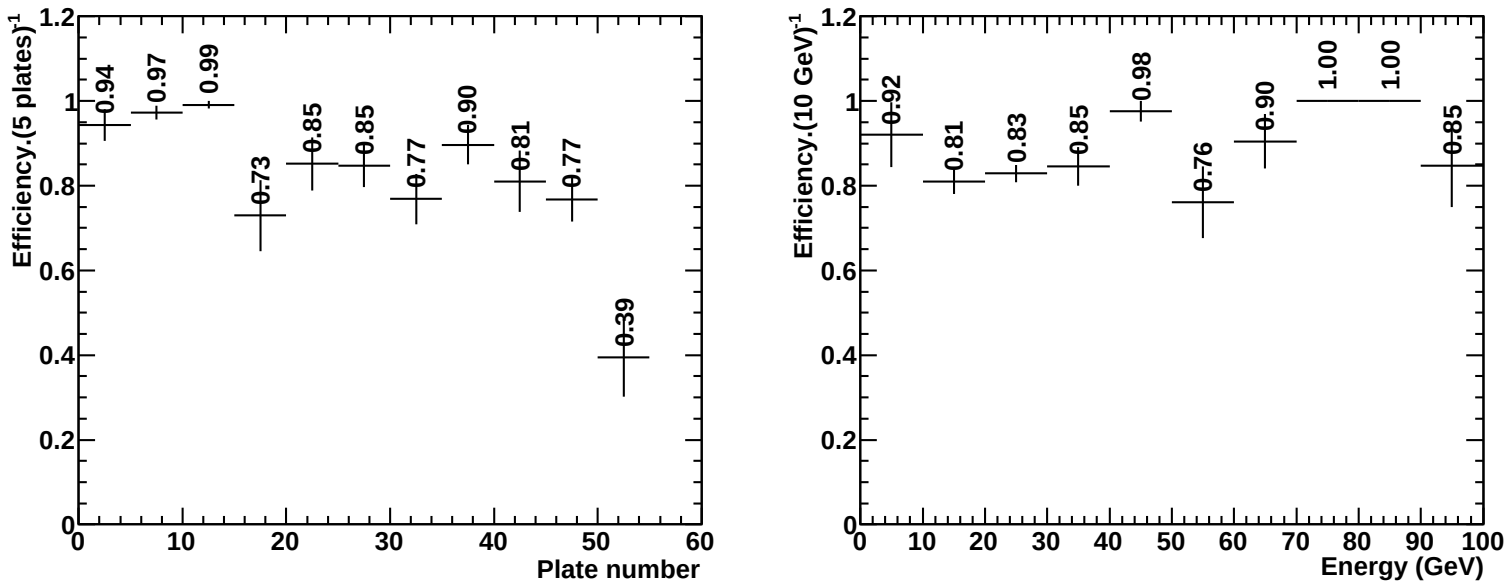


Figure: Electron identification efficiency as a function of the electron MC true energy in the right panel and electron identification efficiency as a function of the MC true vertex depth in the left panel, both obtained with the MC simulation for $\nu_\mu \rightarrow \nu_\tau$ oscillation channel with the τ lepton reconstructed into its electron decay channel.

Performance and efficiency of electron identification

Summary of performance and efficiency of the shower algorithm in electron identification :

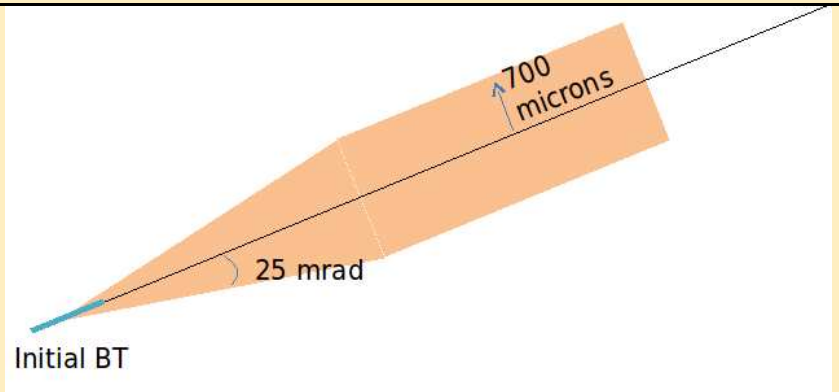
- 10 plates are sufficient to reach an identification efficiency higher than 80% whereas the use of 5 to 10 plates presents an important drop of the efficiency.
- Using MC simulations, the identification effiency with 10 plates have been estimated to 81 $\pm$ 9% for the $\nu_\mu \rightarrow \nu_e$ oscillation channel with a purity of 78%. Systematic uncertainty can be reduce to 4% by using all films available once the electron has been identified.
- Using MC simulations, the identification effiency with 10 plates have been estimated to 87 $\pm$ 13% for the $\nu_\mu \rightarrow \nu_\tau(\tau \rightarrow e)$ channel with a purity of 71%. Systematic uncertainty can be reduce to 7% by using all films available once the electron has been identified.
- Comparison of testbeam data with the MC sample "testbeam-like" shows the difficulty to rely our MC since the instrumental background is missing.

Contents

Energy estimation

The algorithm is based on the same ANN as for the identification algorithm.
However an important discrepancy is on the parameters of the basetrack selection.

Summary of selections specific to energy estimation algorithm in the basetrack collection process.

Variables used for the selections	Values of the selections
Angle difference between the basetrack and the projection of the initiator	$abs(\theta^{3D}_{btk} - \theta^{3D}_{init}) < 0.130\,rd$
The basetrack has to be inside a cone shown on the right	

Strategy to assess performance of the electron energy estimation.

- Estimate the energy estimation with a MC simulation of single eletrons : 10000 electrons produced in an energy continuum from 0 to 30 GeV with a flat spectrum and with incoming angles θ_x and θ_y in the range 0 to 0.6 rad.
- The first step of the analysis has been to define a pure sample of electrons to assess the best performance in energy estimation of this tool.
- Calibration : we release these selections and define a correction to the energy and finally a resolution is given for the OPERA calorimeter with ECC.
- Compare testbeam data with MC simulation to estimate the systematic uncertainty on the energy.
- Check the energy estimation results with MC simulations : 1000 ν_e events coming from the oscillation and 3000 ν_τ events coming from the oscillation with the tau in its electronic decay.

Performance of the energy estimation

Summary of selections to define a pure electron sample.

Variables used for the selections	Values of the selections
Incoming angle of the electron	$abs(\theta^{3D}) < 0.1\,rd$
Contained shower : multiplicity of microtracks	At least 50% of microtracks inside one brick

Showers have been reconstructed over the full volume available for the event.

Performance of the energy estimation

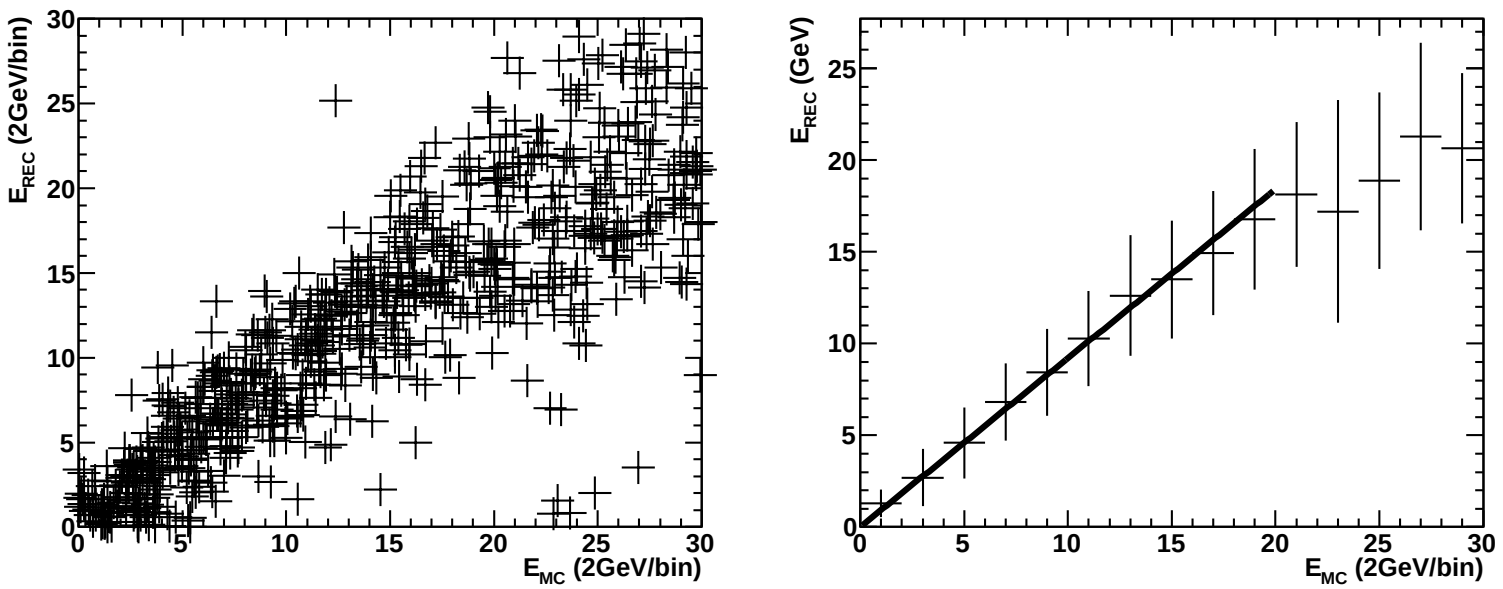


Figure: On the left side a scatter plot of the reconstructed energy as a function of the MC true energy. On the right side the fitted profile of the scatter plot shown on the left. The parametrization showed on the profile gives : $E_{REC} = (0.92 \pm 0.08) * E_{MC}$. The error quoted here is the error on the fit.

The reconstructed energy is well linear up to 20 GeV.

Performance of the energy estimation

$$\text{ratio} = \frac{E_{REC}}{E_{MC}} \qquad \text{residual} = \frac{E_{REC} - E_{MC}}{E_{MC}} = \frac{\Delta E}{E}$$

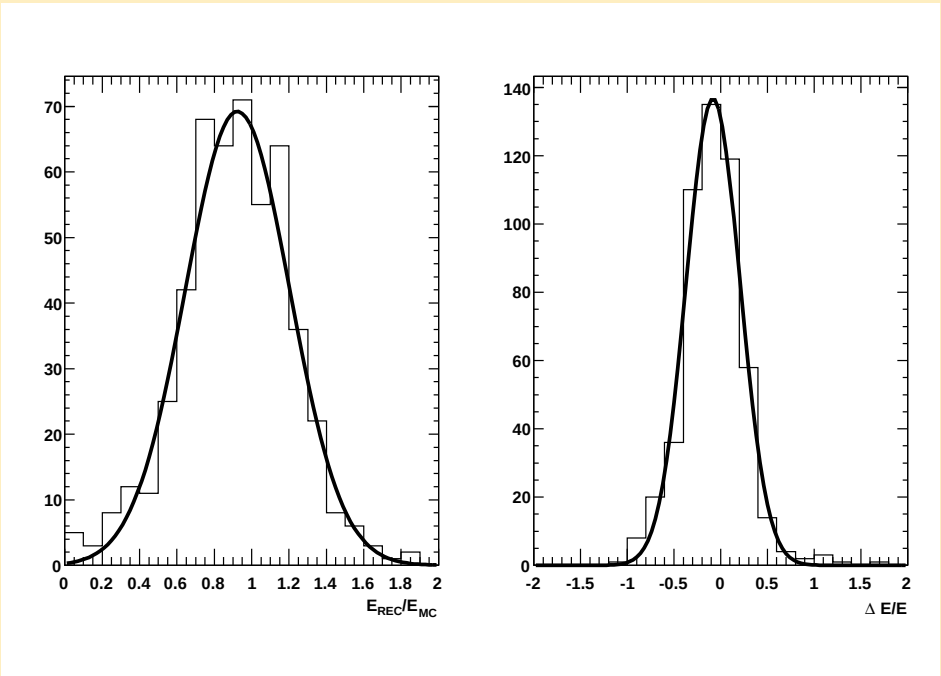
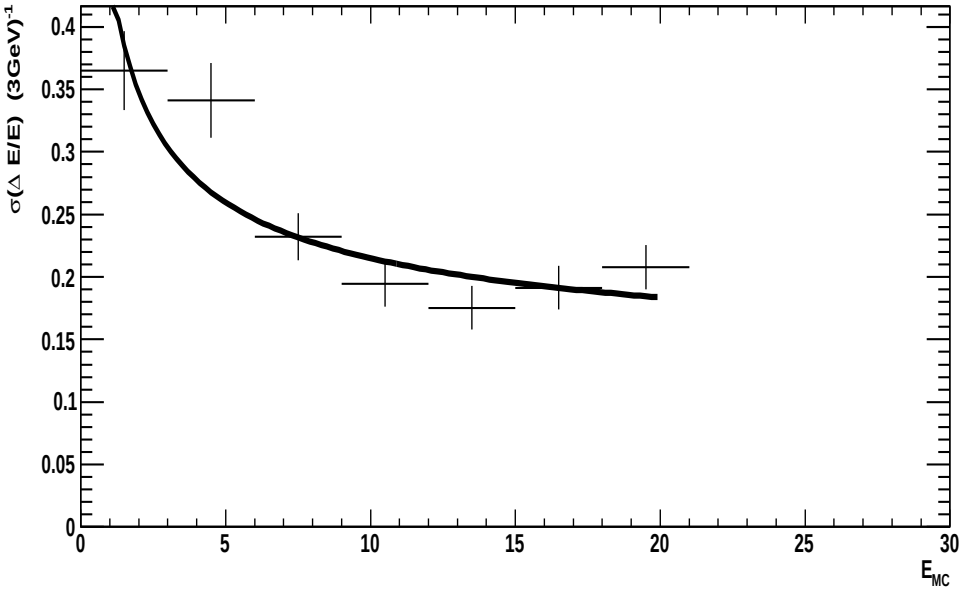


Figure: The gaussian fit of the energy ration gives $\sigma(E_{REC}/E_{MC}) = 0.28 \pm 0.01$; for the residuals, $\sigma(\Delta E/E) = 0.29 \pm 0.01$ and the mean gaussian value is -0.08 ± 0.01.

The energy estimation on the pure sample can achieve a mean resolution of 29%.

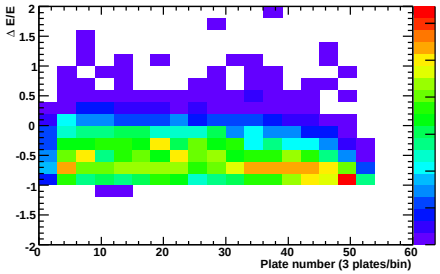
Performance of the energy estimation

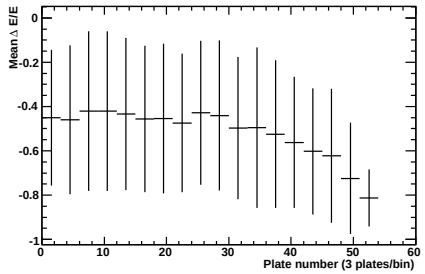
Resolution on the shower energy as a function of the MC true energy.

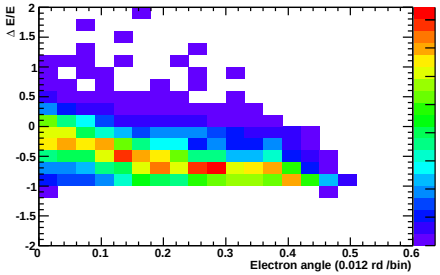


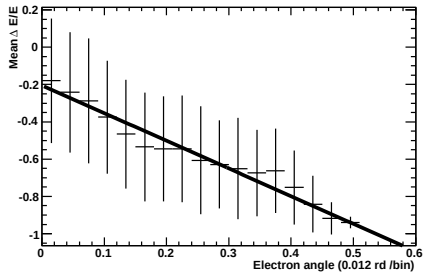
Idealistic case : $\frac{\sigma_E}{E} = \frac{(34 \pm 5)\%}{\sqrt{E}} + (11 \pm 2)\%$

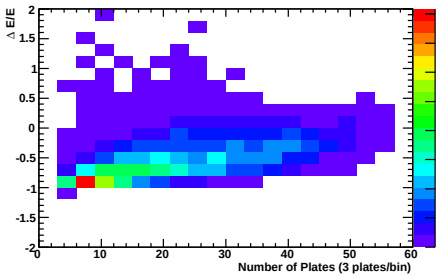
Calibration of the energy estimation

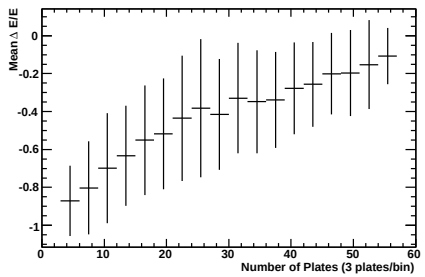








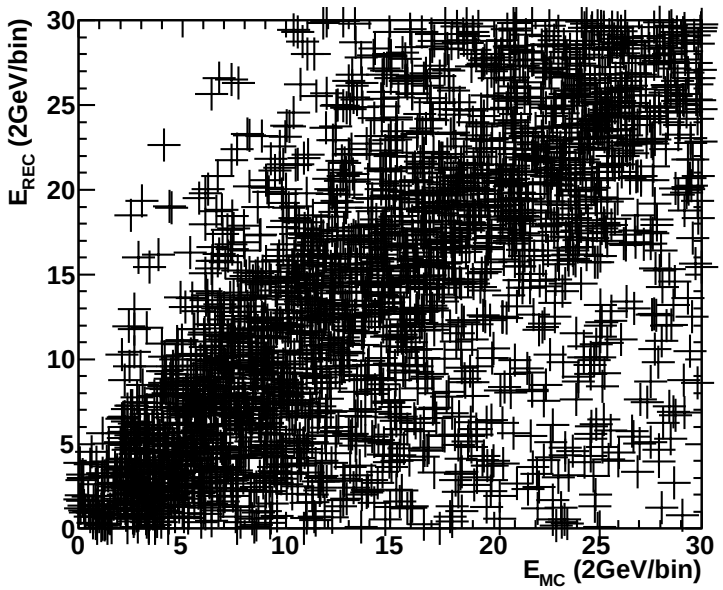


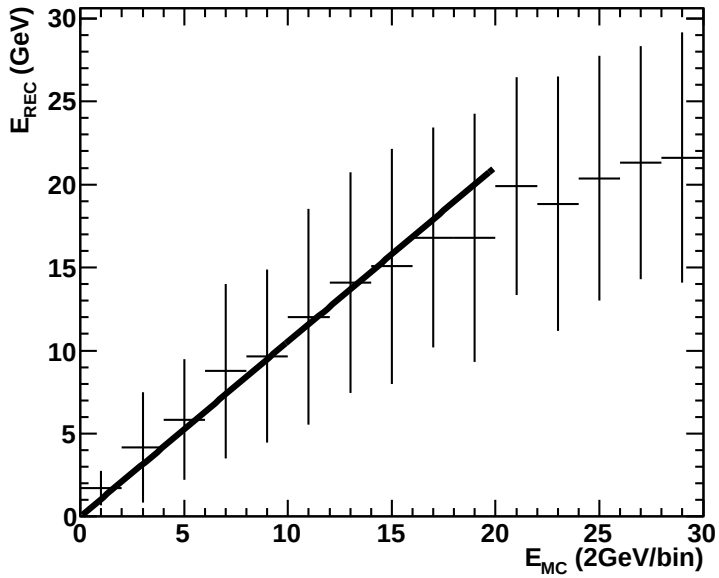


- The energy residuals are stable with the vertex depth up to the plate 40 ; the deviation above this value is visible but not so important up to plate 47.
- The energy residuals vary strongly with the shower axis angle which implies a bias in the energy estimation.
- The shower extension has to be 20 plates at least to reach the stability region in shower extension.

Calibration of the energy estimation

The corrected energy of "contained" showers.

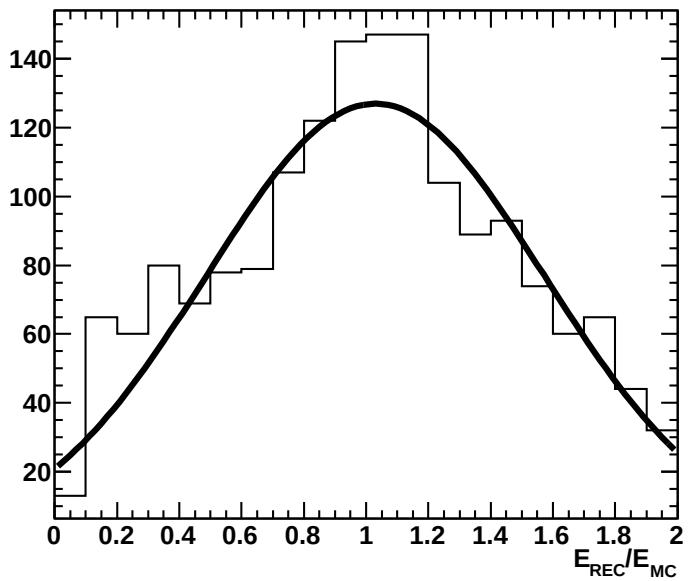


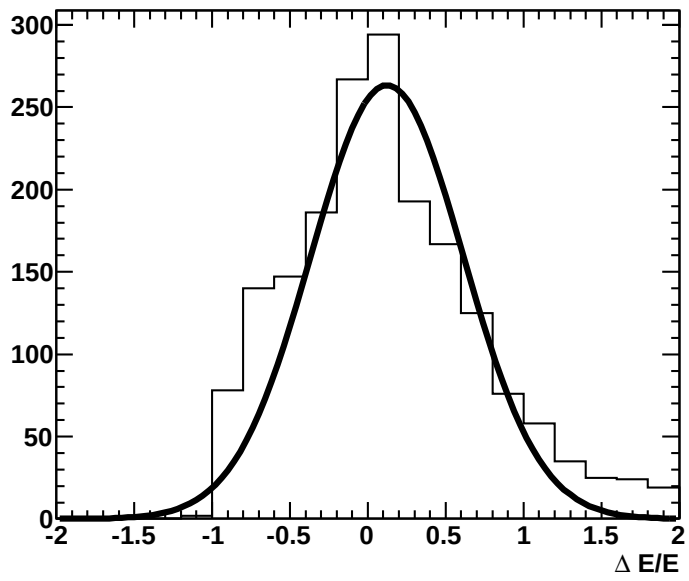


The parametrization showed on the profile gives : $E_{REC} = (1.05 \pm 0.17) * E_{MC}$.

Calibration of the energy estimation

The corrected energy of "contained" showers.

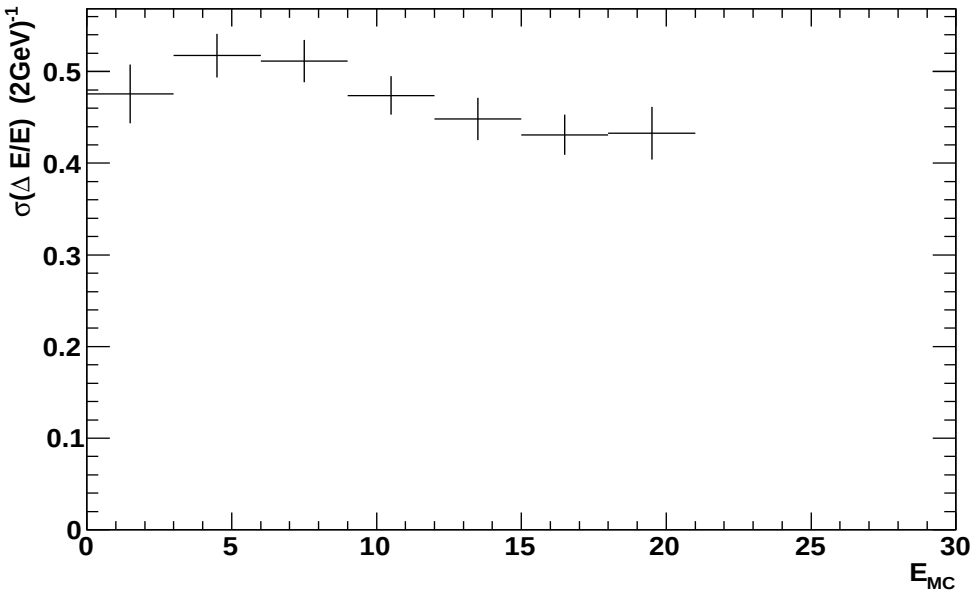




- The gaussian fit of the residuals, $\sigma(\Delta E/E) = 0.49 \pm 0.01$ and the mean gaussian value is 0.12 ± 0.02.
- In this more realistic approach, the correction has improved the reconstructed/MC energy ratio but the resolution on the energy has been enhanced.
- A multivariate correction (shower axis angle, shower extension and vertex depth) would be suited to improve that resolution.

Calibration of the energy estimation

Resolution on the shower energy corrected as a function of the MC true energy.



The final resolution of the OPERA calorimeter is therefore constant with the energy and is about 50%. It does not fit with the usual stochastic behaviour in $1/\sqrt{E}$ because of a too low sampling with 10 X_0 in one brick.

Testbeam data analysis

Ariga's testbeam data analysis

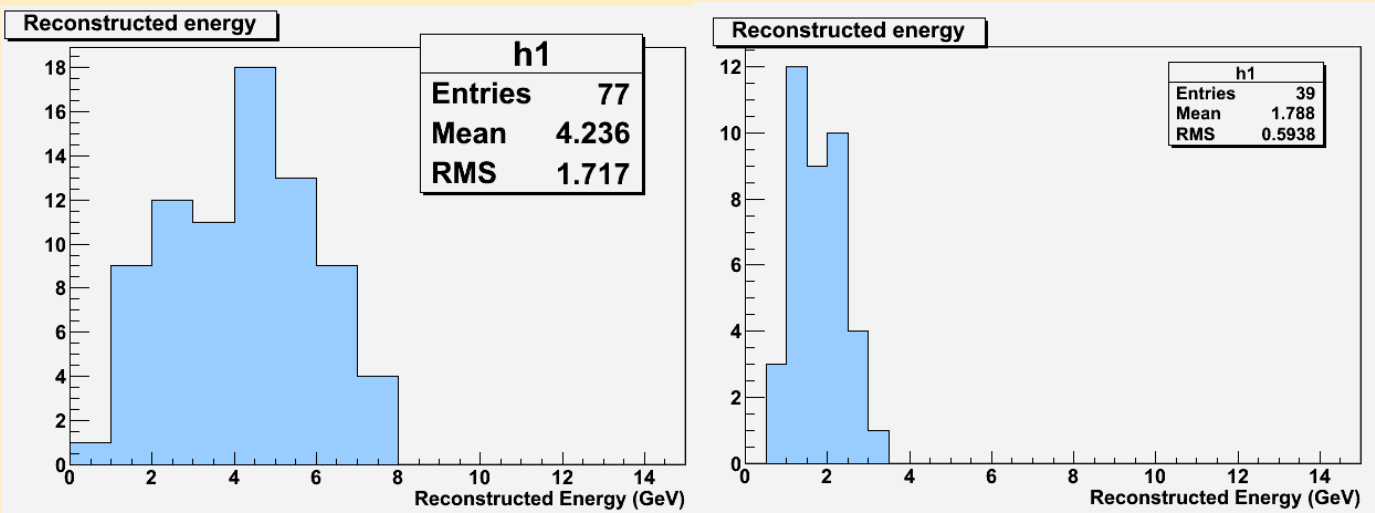
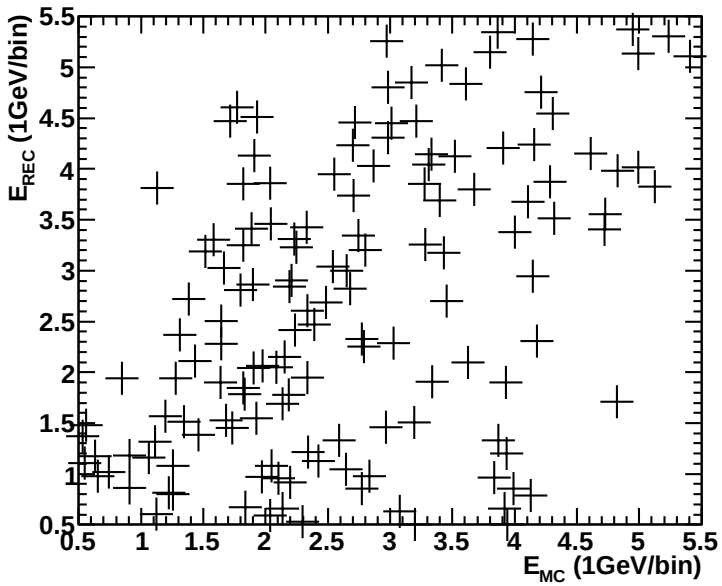
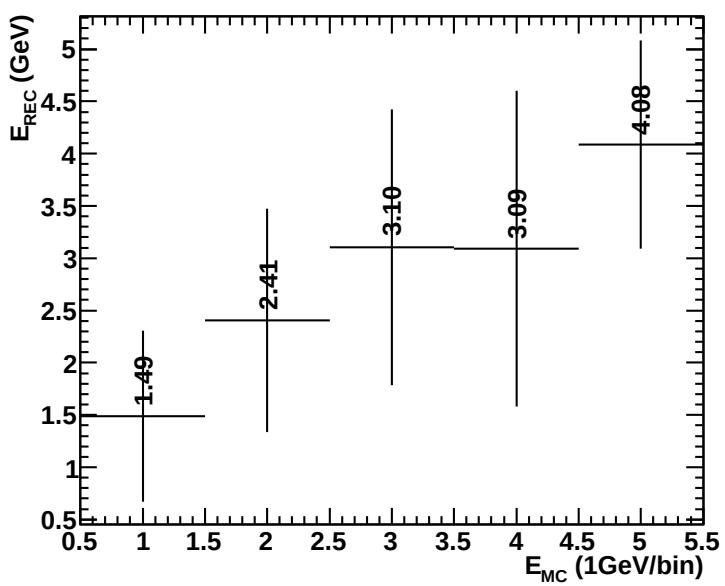


Figure: Reconstructed energy on the selected electrons in testbeam data for 4 GeV (left) and 2 GeV (right).

MC simulation to be compared at testbeam data.

The energy estimation results obtained with the MC simulation of 10 000 single electrons sample for energies between 0.5 and 5.5 GeV and $\theta < 0.1$ rd.





The drop of the bin at 4 GeV is a statistical effect.

Systematic uncertainty on the energy : comparison of MC simulation with the testbeam data.

Summary of the electron energy estimation with testbeam data and MC simulation and the inferred systematic uncertainty.

MC Energy (GeV)	2.4	3.6
Data Energy (GeV)	1.8	4.2
Sys. uncertainty (%)	25	17

The quoted systematic error on the energy estimation is much lower than the resolution.

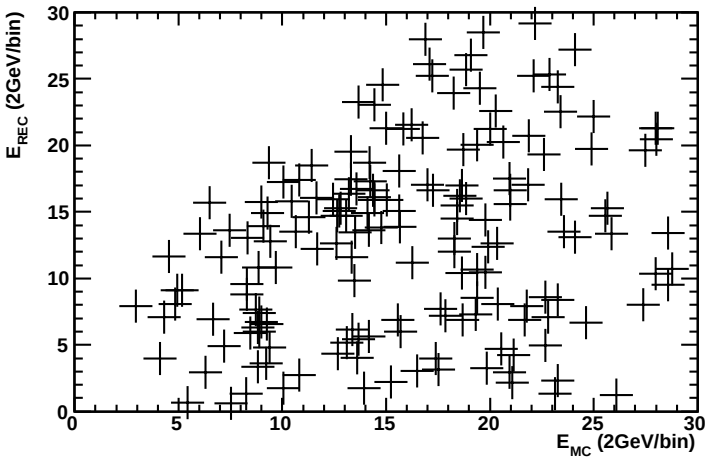
Summary of the performance of energy estimation

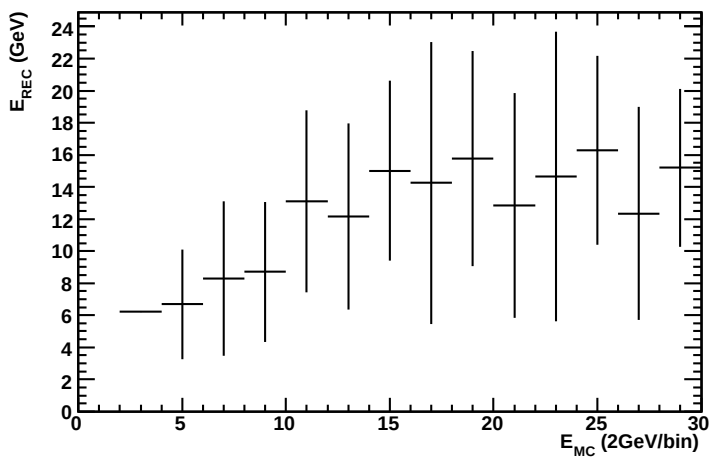
Summary of the performance of the energy estimation of the electron showers :

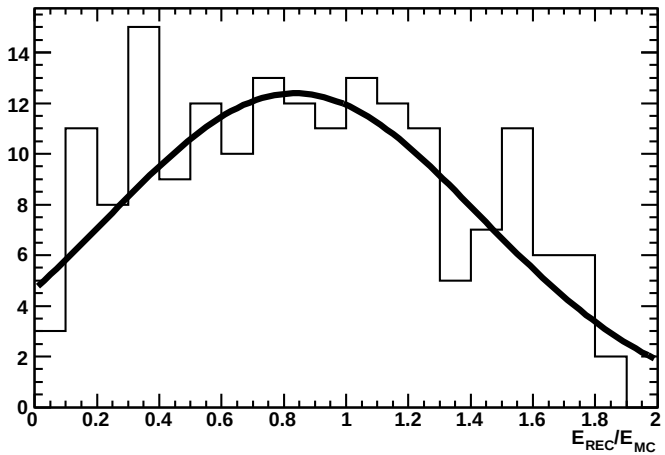
- On a pure sample of electrons : the resolution achieve with the OPERA calorimeter is $\frac{\sigma_E}{E} = \frac{(34 \pm 5)\%}{\sqrt{E}} + (11 \pm 2)\%$.
- The energy is corrected with the shower axis angle and as far as we have at least 20 plates to estimate the energy, the resolution is about 50%.
- The data/MC comparison gives a systematic uncertainty on the energy of 25%.
- Some improvements can be performed : multivariate correction, two-bricks reconstruction but so far the OPERA calorimeter is limited by a low sampling.

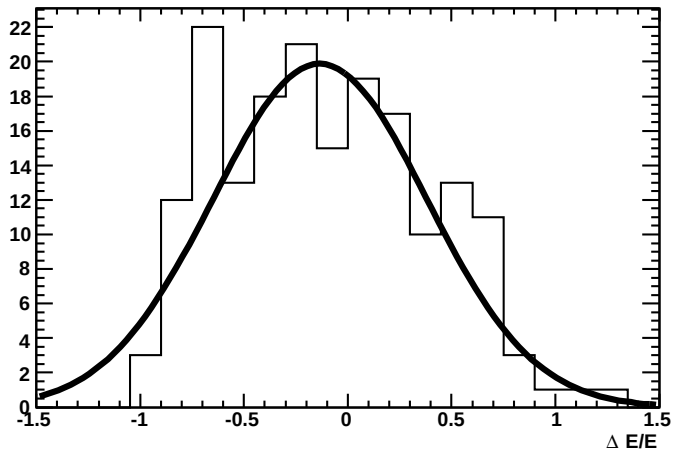
Energy estimation results

Energy estimation of the $\nu_\mu \rightarrow \nu_e$ oscillation sample used in the MC simulation









Conclusion of the shower reconstruction in ECC

- OPERA standard volume (10 plates) is suited to identify the electrons with an efficiency of $81 \pm 7\%$ for the $\nu_\mu \rightarrow \nu_e$ oscillation channel and $87 \pm 15\%$ for the $\nu_\mu \rightarrow \nu_\tau(\tau \rightarrow e)$ channel.
- OPERA standard volume is not sufficient for the energy estimation. About 20 plates are necessary to estimate the energy with a resolution of 50%. A correction to the energy has been defined in order to reduce the bias due to the shower axis angle but have also degraded the resolution from 35% to 50%.

Contents

Oscillated ν_e rate calculation and measurement

Outline of the analysis

- Neutrino event rates computed from the CNGS fluxes, neutrino cross sections and OPERA detector features for all the relevant channel : $\nu_\mu \rightarrow \nu_e$, ν_e prompt, $\tau \rightarrow e$, ν_μ NC and ν_μCC $\xrightarrow{mis}$ ν_μNC. Prompt ν_e QE+RES are also relevant (0.2 ± 0.04 events per 10^{19}p.o.t.) but are still missing in this study.
- Event location efficiency assessment with MC simulation in the OpRelease 4.0 and its systematic uncertainty computed with the muon misidentification in the ED.
- Electron identification by triggering the electron events with CS hints and then identifying with the shower algorithm. The systematic uncertainty of the identification with the shower algorithm has been used.
- Multivariate analysis with 5 kinematical variables : Visible Energy in the TT, Electron energy, missing transverse momentum at primary vertex, phi angle (angle between the hadronic p_T and the p_T of the electron) and the impact parameter instead of using the DSP.

Neutrino event rates computed from the CNGS fluxes, neutrino cross sections and OPERA detector features

	Signal	Backgrounds				
	Oscillated ν_e	$\tau \rightarrow e$	ν_μ^{NC}	prompt ν_e	prompt ν_e QE+RES	$\nu_\mu^{CC} \rightarrow \nu_\mu^{NC}$
N_{evt} exp. in 2008-2009	3.14 $\pm$ 0.48	5.12 $\pm$ 0.52	1100 $\pm$ 111	15.05 $\pm$ 1.06	1.48 $\pm$ 0.11	144 $\pm$ 15
N_{evt} @ 22.5 $\times$ 10^{19}p.o.t.	13.3$\pm$2.0	21.7$\pm$2.2	4670 $\pm$ 473	63.9$\pm$4.5	6.3 $\pm$ 0.5	612 $\pm$ 62

Uncertainties are systematic, calculated from the uncertainties on the beam flux, the cross sections and the non-isoscalar correction.

Event location efficiency assessment with MC simulations

Summary of event location efficiency for each reconstruction step up to the location of the neutrino interaction by the SB procedure.

	oscillated ν_e	$\tau \rightarrow e$	ν_μ^{NC}	prompt ν_e	prompt ν_e QE+RES	$\nu_\mu^{CC} \xrightarrow{\text{mis}} \nu_\mu^{NC}$
$\overline{\text{Muon Id}}$	0.897 ± 0.010	0.852 ± 0.009	0.731 ± 0.008	0.752 ± 0.014	0.943 ± 0.007	0.073 ± 0.005
BF	0.784 ± 0.011	0.727 ± 0.015	0.537 ± 0.009	0.669 ± 0.015	0.841 ± 0.010	0.060 ± 0.025
CS	0.715 ± 0.026	0.612 ± 0.036	0.376 ± 0.016	0.660 ± 0.024	0.807 ± 0.022	0.053 ± 0.007
SB	0.642 ± 0.033	0.524 ± 0.054	0.340 ± 0.018	0.636 ± 0.028	0.730 ± 0.029	0.052 ± 0.007

- The efficiency of each step is cumulative with respect to previous steps.
- The uncertainties quoted are statistical.
- Numbers from the BF up to the location are extrapolated from the 1-brick BF, CS and SB efficiencies since the OpRelease 4.0 cannot process the 2nd brick.

Event location efficiency : Data/MC comparison

Summary of event location in 2008 and 2009 data and comparison with MC simulations.

	2008		2009	
	0μ	1μ	0μ	1μ
Events found by the ED	406	1292	1097	2460
Events located in ECC	169	834	360	1490
DATA location efficiency	0.416 ± 0.024	0.646 ± 0.013	0.328 ± 0.014	0.605 ± 0.010
MC location efficiency with 2 CS tracks	0.340 ± 0.018	0.579 ± 0.029	0.340 ± 0.018	0.579 ± 0.029
MC location efficiency with 1 CS track	-	0.775 ± 0.026	-	0.775 ± 0.026

The MC location efficiency is added for 0μ (ν_μ NC) and the event location efficiency for the 1μ events is computed as for the ν_μ CC events.

- The data/MC differences give the systematic uncertainties.
- The reality of location efficiency is probably in between the 2 last rows.

Event location efficiency : systematic uncertainty

Another way to address event location systematic uncertainty : summary of event location efficiency with or without muon id.

	oscillated ν_e	$\tau \rightarrow e$	ν_μ^{NC}	prompt ν_e	prompt ν_e QE+RES	$\nu_\mu^{CC} \xrightarrow{\text{mis}} \nu_\mu^{NC}$
Standard event location	0.642 ± 0.033	0.524 ± 0.054	0.340 ± 0.018	0.636 ± 0.028	0.730 ± 0.029	-
Simplified event location	0.695 ± 0.031	0.585 ± 0.054	0.430 ± 0.019	0.805 ± 0.024	0.732 ± 0.029	-
Sys. error	0.053	0.061	0.09	0.169	0.002	0.129

The systematic uncertainties are equivalent to the table on the last slide except for the $\nu_\mu^{CC} \xrightarrow{\text{mis}} \nu_\mu^{NC}$ channel. The systematic uncertainty inferred from the data/MC comparison will be used for the latter.

Electron identification efficiency : preliminary study.

Electron shower reconstruction represents a time-consuming computation process.
One has to check if other existing background sources would contribute to this study.

QE+RE interactions and charm $\rightarrow e$ channel rates in the OPERA detector at the event location level.

	charm $\rightarrow$ e	prompt ν_e QE+RES	prompt ν_e	$\tau \rightarrow e$ QE+RES	$\tau \rightarrow e$	osc. ν_e QE+RES	osc. ν_e
ν $(10^{19} pot)^{-1}$	2.2 $\pm$ 0.40	0.28 $\pm$ 0.07	2.8 $\pm$ 0.2	0.066$\pm$ 0.007	0.97 $\pm$ 0.098	0.093$\pm$ 0.014	0.59 $\pm$ 0.09
μ misid	0.052$\pm$ 0.007	-	-	-	-	-	-
Evt loc.	0.524$\pm$ 0.061	0.730$\pm$ 0.15	0.636$\pm$ 0.169	0.374$\pm$ 0.039	0.524$\pm$ 0.061	0.642$\pm$ 0.053	0.642$\pm$ 0.053
Loc. ν $(10^{19} pot)^{-1}$	0.060$\pm$ 0.007	0.20 $\pm$ 0.04	1.78 $\pm$ 0.47	0.025$\pm$ 0.003	0.508$\pm$ 0.059	0.060$\pm$ 0.005	0.379$\pm$ 0.031

- *charm* $\rightarrow e$ is at 10% level of the $\tau \rightarrow e$ background.
- $\tau \rightarrow e$ QE+RES represents about 4% of the $\tau \rightarrow e$ DIS events.
- Oscillated ν_e QE+RES events correspond to 16% of the DIS ones.
- ν_e QE+RES is the only one to be considered.

Electron channels search procedure

This ν_e search process has been systematically applied on the localized events in the 2008-2009 data sample. It consists of a strategy based on two algorithms : the CS shower hint and the electromagnetic shower reconstruction and identification.

- *CS shower hint* : first an algorithm looks for clusters of tracks in the CS corresponding to a shower. An electron hint is validated if the cluster is formed of at least 3 tracks or more and if the position difference and angular difference between the cluster and the other CS tracks are strictly less than $2\,mm$ and $150\,mrad$ respectively.
- *Electron identification* : it is applied on a standard 10 plates volume if the former CS shower hint did not recognize the event as an electron and on the whole brick if the CS shower hint algorithm did validate that the event contains an electron. The result of the shower identification algorithm finally distinguishes events containing an electron from those without an electron.

CS shower hint algorithm triggers events in the inefficiency of the electron identification algorithm in a 10 plates volume.
Moreover, additional pattern recognition algorithms are used to improve the electron identification efficiency. Those are CS to vertex hint, SB shower hint and gamma hint.

Oscillated ν_e rate calculation and measurement

Input variables

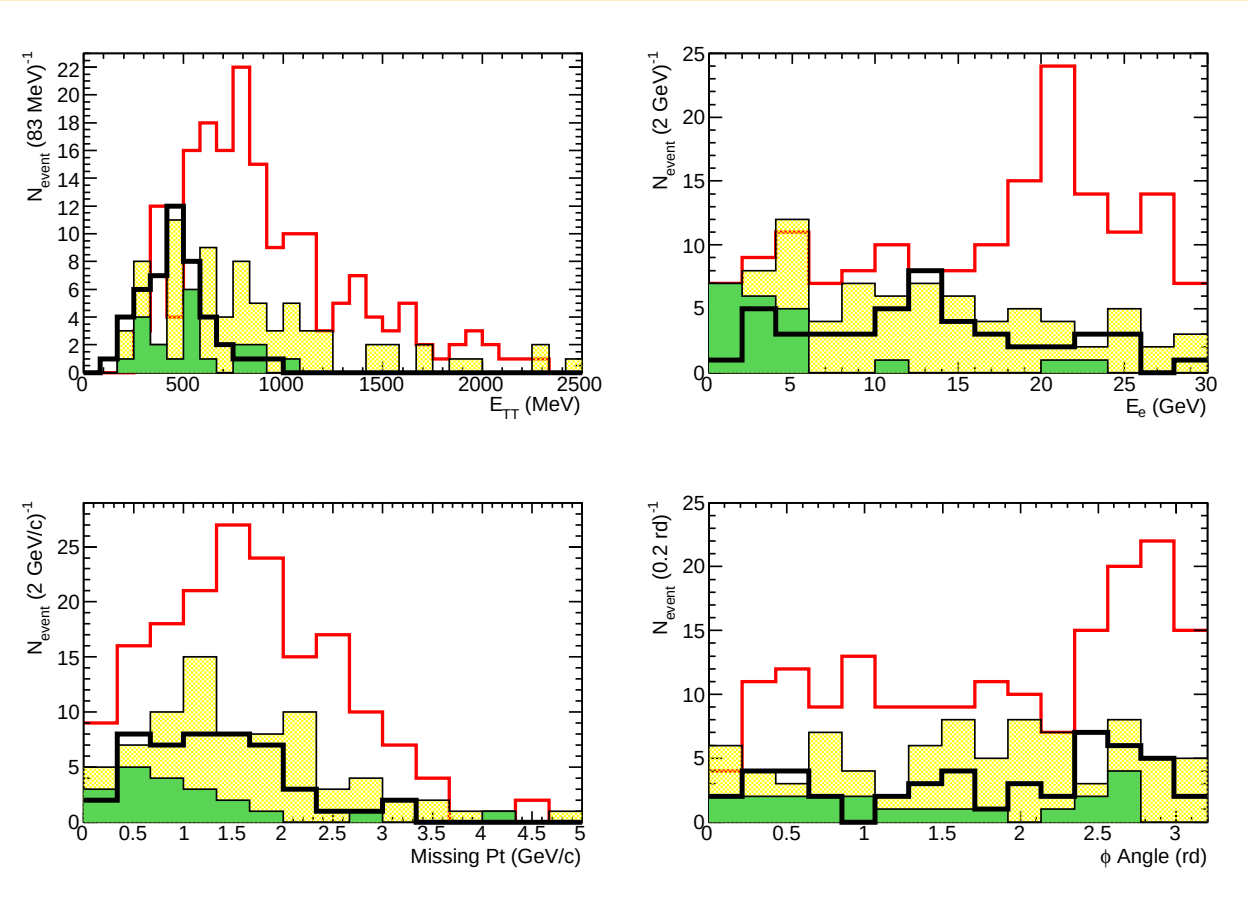


Figure: Distribution of 4 of the input variables of the MVA for the signal and each MC background superimposed. The black thick line represent the signal, i.e. the $\nu_\mu \rightarrow \nu_e(e)$ oscillation channel. The red, yellow and green histograms correspond respectively to the prompt ν_e, the $\nu_\mu \rightarrow \nu_\tau(\tau \rightarrow e)$ and the ν_μ NC channels.

Oscillated ν_e rate calculation and measurement

Input variables

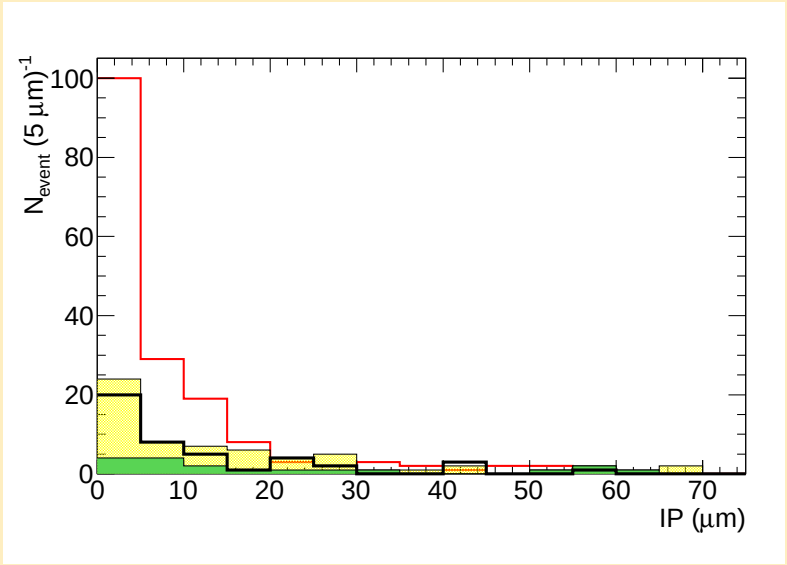


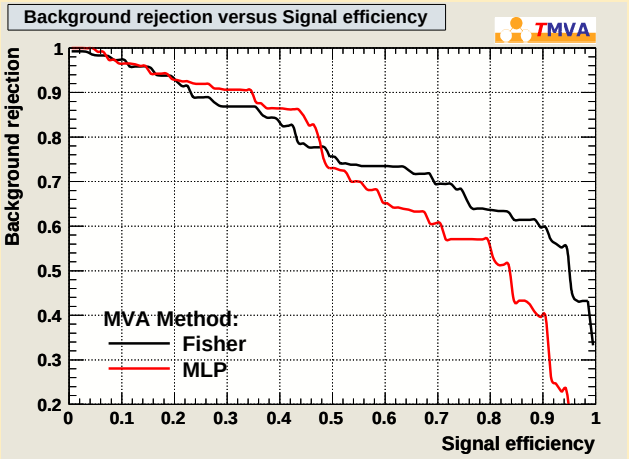
Figure: Distribution of the impact parameter for the signal and each MC background superimposed. The black thick line represent the signal, i.e. the $\nu_\mu \rightarrow \nu_e(e)$ oscillation channel. The red, yellow and green histograms correspond respectively to the prompt ν_e, the $\nu_\mu \rightarrow \nu_\tau(\tau \rightarrow e)$ and the ν_μ NC channels.

- The IP, the missing p_T and the phi angle do not have an important discrimating power. The limit to this study is the definition of the electron track which is not trivial.
- Another problem is the definition of the primary vertex which can be far from the MC true vertex, i.e. inside the shower.

Oscillated ν_e rate calculation and measurement

Input variables distributions feed a multivariate analysis (MVA) using two methods to discriminate background sources : a cut on the Fisher discriminant or an ANN to classify signal and background events.

Multivariate analysis



Efficiency versus purity for the ν_e signal with respect to backgrounds for the two methods Fisher and ANN.

- The ANN method gives better results at high background rejection.
- The Fisher gives better results at high signal efficiency.
- To keep the signal efficiency as high as possible, Fisher method will be used.

Oscillated ν_e rate calculation and measurement

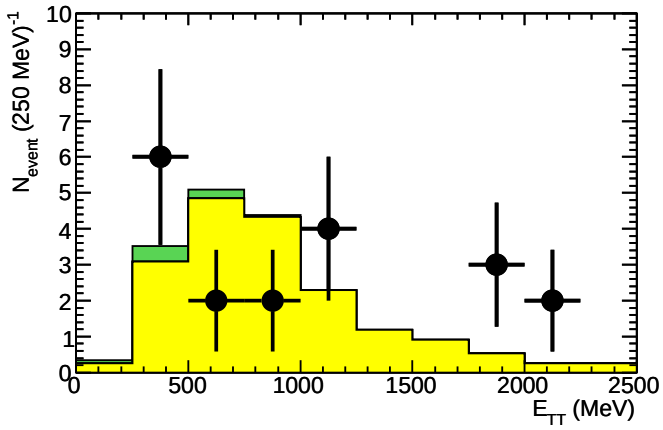
Event rates

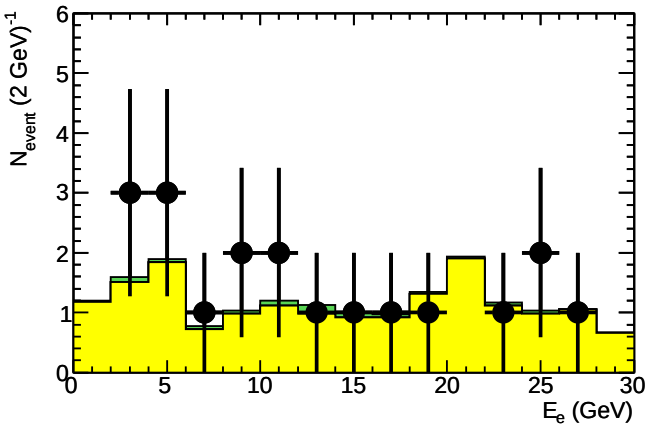
	signal	prompt ν_e	$\tau \rightarrow e$	ν_μ^{NC}	$\nu_\mu^{CC} \rightarrow \nu_\mu^{NC}$
N_{evt} exp. in 2008-2009	1.03 ± 0.23	17.1 ± 5.4	1.1 ± 0.6	6.4 ± 2.1	19.2 ± 5.5
N_{evt} @ 22.5×10^{19}p.o.t.	4.4 ± 1.0	73 ± 23	4.7 ± 2.5	27 ± 9	82 ± 23
Significance in 2008-2009	-	0.24	0.71	0.38	0.23
N_{evt} exp. in 2008-2009 after MVA	0.97 ± 0.22	4.9 ± 1.6	0.50 ± 0.28	1.2 ± 0.4	3.6 ± 1.0
N_{evt} @ 22.5×10^{19}p.o.t. after MVA	4.1 ± 0.9	21 ± 7	2.1 ± 1.2	5.1 ± 1.7	15 ± 4
Significance in 08-09 after MVA	-	0.40	0.80	0.66	0.45

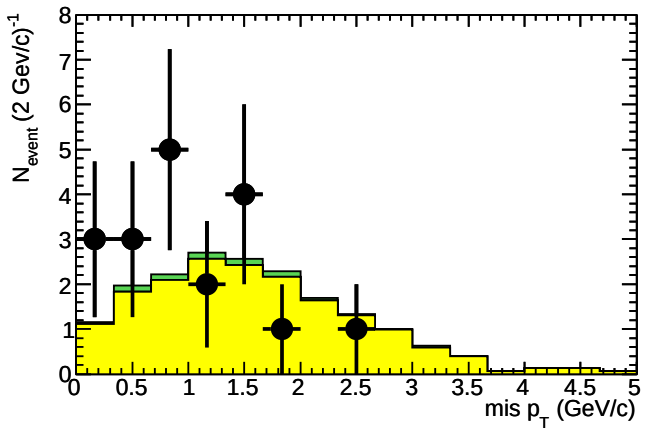
- The MVA reduce all background sources significantly.
- The observation of an oscillation signal is possible with a significance of 0.6σ
- The gamma background sources contribution is equivalent to the prompt ν_e one.

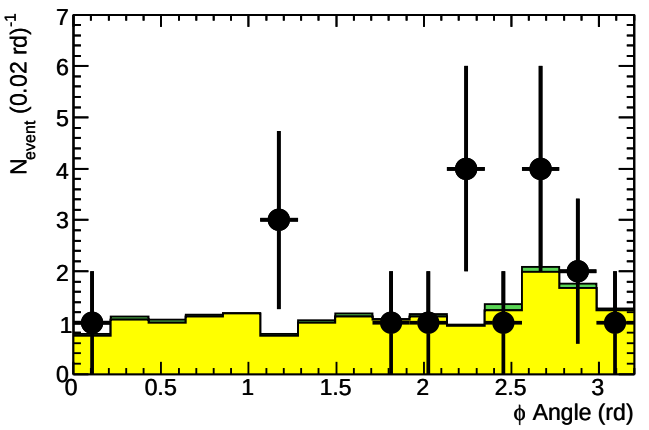
Oscillated ν_e rate calculation and measurement

Comparison of MC with data for 4 of the kinematical variables.



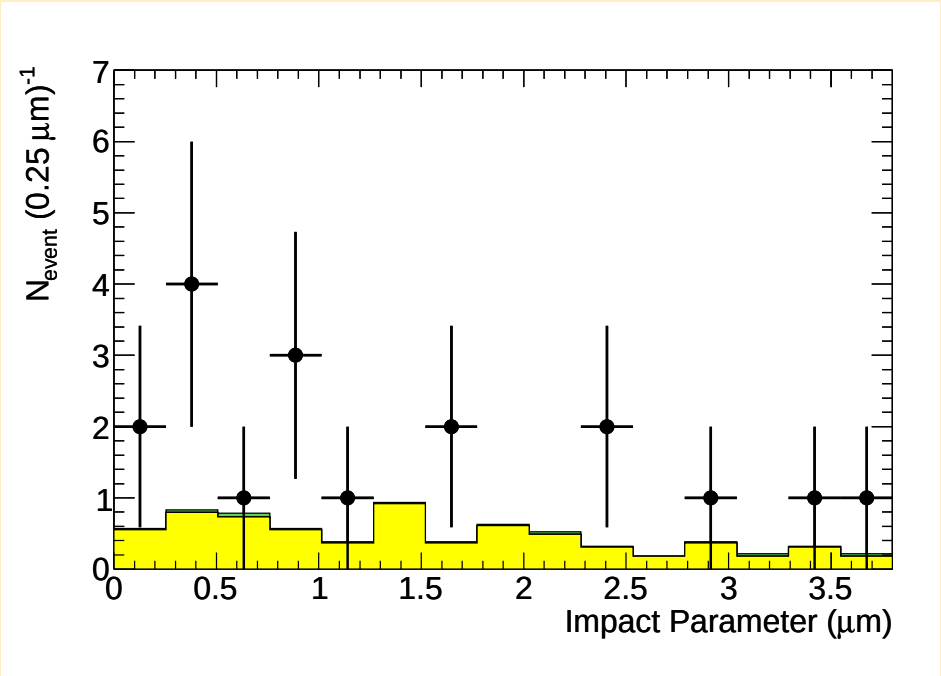






Oscillated ν_e rate calculation and measurement

Comparison of MC with data for the impact parameter.



Comments

- The MC distributions are normalized to data.
- The agreement between MC and data for the considered variables is quite good.

τ_e rate calculation and measurement

Event rates

	$\tau \to e$ signal	oscillated ν_e	prompt ν_e	ν_μ^{NC}	$\nu_\mu^{CC} \to \nu_\mu^{NC}$
N_{evt} exp. in 2008-2009	1.1 ± 0.6	1.03 ± 0.23	17.1 ± 5.4	6.4 ± 2.1	19.2 ± 5.5
N_{evt} @ 22.5×10^{19}p.o.t.	4.7 ± 2.5	4.4 ± 1.0	73 ± 23	27 ± 9	82 ± 23
Significance @ 22.5×10^{19}p.o.t.	-	1.56	0.53	0.83	0.50
N_{evt} @ 22.5×10^{19}p.o.t. after MVA	4.7 ± 2.5	3.7 ± 0.8	66 ± 21	25 ± 8	77 ± 22
Significance @ 22.5×10^{19}p.o.t. after MVA	-	1.62	0.56	0.86	0.52

- The MVA cannot discriminate the background sources from the signal.
- The observation of an oscillation signal is possible with a significance of 0.2σ
- The gamma background sources contribution is equivalent to the prompt ν_e one.

Conclusion

Electromagnetic shower reconstruction

- Identification efficiency and energy estimation have been assessed with their systematic uncertainty.
- The shower algorithm is finally ready to be used in the OPERA analysis.

Oscillated ν_e rate calculation and measurement

- The observation of the $\nu_\mu \rightarrow \nu_e$ oscillation is possible with a significance of 0.6σ evaluated on a full MC simulation.
- The gamma contamination reduction performed on data should be studied on MC simulations.
- The agreement between MC and data for the considered variables is quite good.
- Separate the $\tau \rightarrow e$ signal from the background is a really complicated task.

Prospects

- The shower reconstruction within 2 bricks feature needs a dedicated tool.
- Some other improvements in this analysis could be performed (energy correction, MC OpRelease 4.1).