

ISTITUTO NAZIONALE DI FISICA NUCLEARE

Laboratori Nazionali di Frascati

INFN-15-04/LNF 7th April 2015

VIP Data Analysis

Report

Dorel Pietreanu, Laura Sperandio ¹⁾INFN-Laboratori Nazionali di Frascati Via E. Fermi 40, Frascati, Italy

Abstract

In this report the data analyses for the VIP experiment, which took data at the LNGS-INFN laboratory in the period 2006-2010, is presented in detail, together with the final result on the limit of the Pauli Exclusion Principle (PEP) probability for electrons.

The data analysis is performed by using the ROOT software, in the first part without the refined charge transport correction, while in the second part with an optimized procedure for the charge transport correction for the Charged-Coupled Devices (CCDs).

After introducing the structure of the data files, the cuts made to get the files available for the refined analysis are described.

All the steps to obtain the final histograms and the fit results are presented in detail, together with the calculations for the identification of the region of interest and of the limit for the PEP violation probability.

Pubblicato da SIDS–Pubblicazioni Laboratori Nazionali di Frascati

Index

1. Without charge transport correction

- 1.1 VIP data files
- 1.2 VIP data structure
- 1.3 VIP data analysis
- 1.3.1 Histograms
- 1.3.2 Histogram analysis
 - 1.3.2.a) Initial fit
 - 1.3.2.b) Gaussian fit of the Cu K_a and K_b lines
- 1.3.3 Energy re binning and sum of the histograms
 1.3.3.a) Rebinning the channel histograms
 1.3.3.b) Sums of the energy histograms
 1.3.4 Histogram subtraction and β²/2 calculation
 1.3.4.a) Region of interest
 - 1.3.4.b) Experimental $\beta^2/2$ calculation
 - 1.3.4.c) Residual peak
- 1.4 Spectral line identification
- 2. Charge transport correction

VIP Data Analysis Report

1.1 VIP data files

Two kind of data were acquired in VIP experiment: background data and signal data. The signal data were taken when a current of intensity I = 40 A passed through the copper target, while the background data were taken without current.

The periods for each type of data are:

- Background (I = 0A) data: 112602 files = 536010 min;
- Signal (I = 40A) data: 81185 files = 405925 min.

Each file has a duration of 5 minutes. Part of the data were invalidated by power losses, defective pixels, discontinuities in the data taking, etc., and after selecting all valid data, the effective data size is:

- Background (I = 0A) data: 83118 files = 415590 min;
- Signal (I = 40A) data: 79654 files = 398270 min.

1.2 VIP data structure

VIP raw files are the images coming from the 16 CCDs (1248x1152 pixels) of the VIP detector. Each hit pixel records an ADC value of the deposited charge. Using a pre-analysis program, the VIP ntuples are generated from row data as HBOOK files (fortran compatible structures) which are to be summed. In this analysis a bash script was used to transform and add these files in a more compact ROOT based structure. The resulting sets of 16 ROOT files (one file for each CCD and one set for each hard disk of data), contains two trees: the **header** and **event data**. The entries of these trees are different and apparently hard to correlated, but their synchronization has been performed using the

variable events from the data tree.

The variables map is shown in Fig. 1; here the most important of them are described:

- ccd_id \rightarrow identification number of the CCD
- **file_id** → identification number of the file
- **event_adc** \rightarrow content of each pixel
- defect \rightarrow defective pixel, based on the number of counts in that pixel
- grade → indicates if single pixel, double pixel or single + double pixel are considered in the analysis
- **noise_mean** and **noise_deviation** \rightarrow are referred to the noise peak in the spectra
- **x_event** and **y_event** \rightarrow are the X and Y coordinate of the pixel
- **nbr1....24_adc** \rightarrow for each fixed pixel, it indicates the number of the first and second neighbor.



Fig. 1 VIP ntuple – data variables

1.3 VIP data analysis

1.3.1 Histograms

Charge deposited in each pixel was acquired using a 12 bit ADC. The ADC assigns a digital value for each pixel according to its voltage amplitude, as result, the recorded values can be represented using a 4096 channels histogram. Fig. 2 shows the recorded spectra coming from a CCD (eg., ccd2, hardisk3, from now on used as example in this note). We observed that no specific line can be observed in the spectra. The following cuts were applied to obtain the histograms suitable for the analysis.

Files cut. One of the cuts was applied to eliminate the files with anomalous number of entries. This modification was caused mostly by the failure of the cryogenic and air conditioning systems reflecting thus a strong instability of the CCDs with temperature, underlined by the plots in Fig. 3a and Fig. 3b. All values for this cut are shown in table 1, Annex A.



Spectrum Full Hard3 CCD2

Fig. 2 ADC spectrum no cuts



Fig. 3a. Total number of entries for a hard disk of data



Fig. 3b. Total number of entries after the file cut

Some files inside of some hard disks show a high number of counts (above average) due mainly to a temporary malfunctioning of the detectors, subjected to temperature variations (see Fig. 3a and Fig. 3b). To have a good subtraction of the signal-background data, the cut was applied until the ratio **number of entries / number of files** was almost constant for each hard disk (Fig. 4). If this ratio would have different values for different hard disk, fake signals would be obtained in the final spectra after background normalization and subtraction.



Fig. 4 Ratio Entries / Files vs hard disk number for VIP data

Defective pixels. Some of the pixels are fire without being hit by any X-Ray - these are the so called *defective pixels* and they can be identified only by the statistics, after adding the data from a set (one hard disk). In this analysis we take set the defect threshold at 10, i.e., if in a set of data (a hard disk set) a pixel fires more than 10 times, it is considered defective. Fig 5 shows maps of the defective pixels, in two CCDs: we see that CCD1 has many defective pixels.





d. CCD2 zoom



Grade. Another cut was performed on the "nature" of the X-Ray hit. Three type of data were recorded:

• single pixel, grade = 0 (Fig.6b);

- double pixel , grade = 1 (Fig.6c);
- single + double pixel grade ≥ 0 (Fig. 6a).

Due to the very low statistics in the analysis, single + double pixel data was considered.



a. ADC Spectrum obtained with cuts for single & double pixel



9

1.3.2 Histogram analysis

The channels data histograms are obtained using all the cuts described above. From these histograms, using a fit calibration procedure, the aligned energy histograms are generated. Because of the missing calibration data, the *in situ* K_{α} and K_{β} X-rays transitions of the Cu29 have to be used. Having such a low statistics (Fig. 7), the fit results are very sensitive and the histograms required an adequate rebinning (was found to be 8 – see Fig. 8) and fit algorithm.



Fig. 7 CCD 2 histogram - 4096 channels

1.3.2.a) Initial fit

For the background a linear fitting with a polynomial *Pol1* function (1) was performed. The peak regions were excluded from the linear fit in order to obtain the parameters with the best χ squared

(Fig. 8).

$$f_{background} = P0 + x \times P1 \tag{1}$$



Fig. 8 Background linear fit with eq. (1)

Next step was to identify the height of the K_{a1} line using a *peak searcher*. Once it was found, then the mean position is determined and sigma of K_{a1} is roughly calculated using the relation between FWHM and σ : FWHM = 2.35 × σ . These are the three parameters for the Gaussian function of Cu29 line K_{a1} . Then, using a combination of **polynomial + Gaussian + Gaussian + Gaussian** function (Fig. 9) with six free parameters, the background and the K_{a1} , K_{a2} , K_{a} lines of the Cu29 were fitted.

More specifically, the fit parameters are:

- background slope and offset; the starting values are determined from the background fit (2 parameters)
- peak height, mean and sigma of the first Gaussian K_{a1} ; the starting values are set as explained above

• mean of the third Gaussian K_{μ} ; equation (3) is used to set the starting value.

All three Gaussian peaks have the same sigma, which is determined by the detector resolution. The K_{a1} and

 K_{a2} peaks have a separation of about 20 eV, and they contribute to produce asymmetric tails in the K_a peak.

Concerning the energy and the relative intensity of the Cu29 lines the values used in the fit are:

- $K_{a1} \rightarrow E = 8027 \text{ eV} (100)$
- $K_{a2} \rightarrow E = 8047 \text{ eV} (51).$
- $K_{a2} \rightarrow E = 8905.3 \text{ eV} (17).$

$$f_{init}(x) = P_0 + P_1 x + H_{K_{\alpha_1}} \exp\left(-\frac{\left(x - \mu_{K_{\alpha_1}}\right)^2}{2\sigma_{K_{\alpha_1}}^2}\right) + 0.51 H_{K_{\alpha_1}} \exp\left(-\frac{\left(x - \left(\mu_{K_{\alpha_1}} - \frac{\mu_{K_{\beta}} - \mu_{K_{\alpha_1}}}{E_{K_{\beta}} - E_{K_{\alpha_1}}}\left(E_{K_{\alpha_1}} - E_{K_{\alpha_2}}\right)\right)\right)^2\right) + 0.17 H_{K_{\alpha_1}} \exp\left(-\frac{\left(x - \mu_{K_{\beta}}\right)^2}{2\sigma_{K_{\alpha_1}}^2}\right)$$

$$(2)$$



Fig. 9 Initial fitting function

(3)
$$\mu_{K_{\beta}} = \mu_{K_{\alpha_{1}}} + \frac{\mu_{K_{\alpha_{1}}}}{E_{K_{\alpha_{1}}}} \left(E_{K_{\beta}} - E_{K_{\alpha_{1}}} \right)$$



Fig. 10 Initial fit parameters. Sigma K_{α} = Sigma K_{β}

1.3.2.b) Gaussian fit of the Cu K_a and K_b lines

The main purpose of this fit is to determine with a good precision the mean of the K_{α} and K_{β} lines of the Cu29, in order to have a good energy calibration and alignment of the histograms. Once the initial fit provide the parameters, another fit was performed, this time having σ_{β} as free parameter, and using as start value the one of σ_{α} obtained from initial fit and using the energy dependent formula (6). The fit function (4) with the seven free parameters and fit results are shown in Fig. 11.

The fitting function with energy dependent parameterization for σ (5) has been also tried, but because of the very low statistics, the obtained fit results (overall) were worse than the previous parameterization. The total number of free parameters remains 7 (Noise and Fano where used instead of sigma K_a and sigma K_p) but some of the fits were not convergent and non typical values for Fano factor were obtained.

$$f_{fit}(x) = P_0 + P_1 x + H_{K_{\alpha_1}} \exp\left(-\frac{\left(x - \mu_{K_{\alpha_1}}\right)^2}{2\sigma_{K_{\alpha_1}}^2}\right) + 0.51 H_{K_{\alpha_1}} \exp\left(-\frac{\left(x - \left(\mu_{K_{\alpha_1}} - \frac{\mu_{K_{\beta}} - \mu_{K_{\alpha_1}}}{E_{K_{\beta}} - E_{K_{\alpha_1}}}\left(E_{K_{\alpha_1}} - E_{K_{\alpha_2}}\right)\right)\right)^2\right) + 0.17 H_{K_{\alpha_1}} \exp\left(-\frac{\left(x - \mu_{K_{\beta}}\right)^2}{2\sigma_{K_{\beta}}^2}\right)$$
(4)



Fig. 11 Fitting function and parameters, having $\sigma\,K_{\scriptscriptstyle F}$ as free parameter

$$f(x) = P_0 + P_1 x + H_{K_{\alpha_1}} \exp\left(-\frac{\left(x - \mu_{K_{\alpha_1}}\right)^2}{2\sigma_{K_{\alpha_1}}^2}\right) + 0.51 H_{K_{\alpha_1}} \exp\left(-\frac{\left(x - \mu_{K_{\alpha_2}}\right)^2}{2\sigma_{K_{\alpha_2}}^2}\right) + 0.17 H_{K_{\alpha_1}} \exp\left(-\frac{\left(x - \mu_{K_{\beta}}\right)^2}{2\sigma_{K_{\beta}}^2}\right)$$
(5)

with:

$$\sigma_{K_x} = \sqrt{\left(\frac{Noise}{2\sqrt{2\ln 2}}\right)^2 + \epsilon FanoE_{K_x}} \tag{6}$$

where ε is the average energy to create a charge carrier pair in the detector (3.8eV).

1.3.3 Energy re binning and sum of the histograms

1.3.3.a) Rebinning the channel histograms

This analysis requires the sum of the data histograms for both, signal and background events. For the alignment of these histograms the energy rebinning parameters were calculated using the mean of the K_{α} and K_{β} lines of Cu29 (Fig. 12), previously determined by the fit procedure. Each set of data (one hardisk data) has its own set of energy rebinning parameters, used to generate the 4096 channels energy histograms with 6eV/bin (Fig.13) The formula used for rebinning was:

$$Bin_{energy} = P_0 + P_1 \times Bin_{channel} \tag{7}$$



Rebin parameters of fcthardisk3ccdgradef2

Fig. 12 Parameters used for energy rebinning of the ADC spectra

In Fig. 12, the large statistical uncertainty of the second point (K_pmean) is caused by the poor statistics.

XRay Energy hardisk3ccdgradef2



Fig. 13 Energy spectra CCD2

1.3.3.b) Sums of the energy histograms

Two kind of sums were made:

- sum of each CCD for both, *current* and *no current* data;
- sum of the data of each hard disk.

The results are energy histograms, which were fitted in order to obtain the energy resolution for each set of data (hardisk/CCDs). This time, the initial function was parameterized with 4 free parameters (2pol, K_{α} height and K_{α} σ) and the final function with 5 free parameters (2pol, K_{α} height , σK_{α} and σK_{β})

The fit algorithm is the same with the one used to obtain rebin parameters:

- linear fit of the background excluding the peak regions pol1;
- initial fit *pol1*+gauss+gauss+gauss, same sigma for K_α and K_β lines;
- final fit *-poll+gauss+gauss+gauss*, having sigma K_{μ} as free parameter.

Partial results of these fits are shown in Fig. 15, Fig. 16, and Fig. 17 and full results of resolution are plotted in Fig. 18 – Fig.21.



Fig. 14 Background fit for the CCD2, current histogram



Fig. 15 Initial fit for the CCD2, current histogram



Fig. 16 Final fit for the CCD2, current histogram



Fig. 17 Fit for hardisk 3 histogram



Fig. 18 Resolutions of K_{α} and K_{β} lines, for each hardisk, background data



Energy resolution current data

Fig. 19 Resolutions of K_α and K_β lines, for each hardisk, signal data



Fig. 20 Resolutions of K_α and K_β lines, for each CCD, background data



Energy resolution current data

Fig. 21 Resolutions of K_{α} and K_{β} lines, for each CCD, signal data

1.3.4 Histogram subtraction and $\beta^2/2$ calculation 1.3.4.a) Region of interest

VIP is searching for anomalous (violating) transition in cooper atoms. A theoretical value of the X-Ray violating transition was estimated using a Dirack Fock method and a software algorithm for muonic atoms [1] [2]. This value defines a *region of interest RoI*, in which experimental searches are performed. Of course, this RoI is strongly dependent with the resolution of the detector at that energy, $(\Gamma = FWHM = (2.354 \cdot \sigma) \text{ eV})$:

$$RoI = \left[7729 \pm \frac{(\Gamma + 10)}{2}\right] eV$$
⁽⁸⁾

The final VIP data are summarized by the two histograms (*signal* and *background* or *current* and *no current* data) which are to be subtracted in order to identify possible signals in the region of interest (Fig.22) defined by the largest Gamma in the signal and background histograms (Fig.22 and Fig.23). Considering the FWHM of the signal histogram, formula (8) will give RoI=[7536, 7922].



Fig. 22 Signal (I = 40A) histogram



Fig. 23 Background (I = 0A) histogram

To get the best fit results the final histogram have been fitted with different binning values. It was noticed (Fig. 24) that the binning does not have a strong influence on σ but on the values of χ^2 .



Fig. 24 Signal and background histogram rebinned



Fig. 25 Signal and scaled background histograms

Before the *signal-background subtraction*, the background histogram was scaled with the files ratio factor (in our case 79654/83118 = 0.95832431), as consequence, the normalization was made with respect to time (each file having an acquisition time of 5 minutes). Fig. 25 shows the spectra of the signal and scaled background data and the excellent correspondence of the shape of the two plots highlights the goodness of the file cut procedure.

1.3.4.b) Experimental $\beta^2/2$ calculation

To determine the experimental limit on $\beta^2/2$ from our data, we used the same arguments of Ramberg and Snow [3] [4], comparing thus the experimental results. The $\beta^2/2$ parameter is related here to the probability that an electron violates PEP. The total number of electrons passing through the conductor, considered "new " or "fresh" for this conductor, is given by:

$$N_{new} = \frac{1}{e} \int I dt \tag{9}$$

where *e* is the electron electric charge, *I* is the current intensity and Δt represents the time duration of the measurement with current on.

Each new electron will undergo a large number of scattering processes on the atoms of the copper lattice. The minimum number of these internal scattering processes per electron, defined as N_{int} , is of order D/μ , where D is the length of the copper electrode (8.8 cm in our case) and μ is the mean free path of electrons in copper. The latter parameter is obtained from the resistivity of the metal. It was assumed that the capture probability (aside from the factor $\approx \beta^2/2$) is greater than 1/10 of the scattering probability.

The acceptance of the 16 CCD detectors and the probability that an X-ray of about 7.6 keV, the energy of the possible anomalous transition generated in the copper target, is not absorbed inside the copper itself, were evaluated by a Monte Carlo simulation of the VIP setup, based on GEANT 3.21. This probability turns out to be 2.1%. Moreover, a CCD efficiency equal to 48% for a 7.6 keV X-ray was considered. All these factors built up the so-called *geometric factor* - **GF**(\sim 1%).

The number of X-rays generated in the PEP violating transition, ΔN_X , is then related to the $\beta^2/2$ parameter by:

$$\Delta N_x \ge \frac{1}{2} \beta^2 N_{new} \frac{1}{10} N_{int} \times GF \tag{10}$$

$$\Delta N_x \ge \beta^2 \frac{1}{20e\mu} \left(\Sigma I \Delta t \right) \times GF \tag{11}$$

Using the values for $\sum I \Delta t = 34.824 \times 10^6$ C, D = 8.8 cm, $\mu = 3.9 \times 10^{-6}$ cm, $e = 1.602 \times 10^{-19}$ C, results

$$\frac{\beta^2}{2} \le \frac{3 \times \Delta N_x}{FileNo \times 300 \times 40} \times 7.1 \times 10^{-23}$$
⁽¹²⁾

Formula (12) couples the violation parameter to the number of X-rays found in the region of interest.

The difference of events between the measurements with and without current, into the RoI, is $\Delta NX = 179 \pm 182$. Taking as a limit of observation three standard deviations, we get for the PEP violating parameter:

$$\beta^2/2 \le 4.05 \times 10^{-29}$$
 at 99.7 C.L. (13)

This can be interpret as a limit on the probability of PEP violating interactions between external electrons and copper atoms.



Fig. 26 Signal and scaled background histogram subtraction



Final VIP subtracted spectra

Fig. 27 Region of interest

1.3.4.c) Residual peak

When the current flows through the conductor, the CCDs are heated by a few Celsius degrees. This is highlighted by the noise recorded data (Fig. 28 b)), and the differences in resolution between the signal and background data. The waving trend (Fig. 28 a)) recorded by the slow control is reproducing the air conditioning cycle in the VIP barrack.



Fig. 28 a) Variation of the target (left) and CCDs (right) temperature during DAQ



Fig. 28 b) Noise recorded by ADC for a VIP data set

Different resolutions for the signal and background data will cause residual peak most visible at 8 keV. Such different resolutions are due to the fact that when the signal is detected, the target is heated and detectors working at a higher temperature. A difference, even a few eV, between the two resolutions is more evident in correspondence of the 8 keV, peak where the number of counts focuses.

The residual peak was studied by the signal and background histograms generated from the fit functions with their errors taken from the fit covariance matrix (Fig. 29) or calculated with Poisson rule





Fig. 29 Subtraction of the two fit functions with fit errors



Fig. 30 Subtraction of the two fit functions - Poisson errors

Their subtraction reveals the 8 keV residual peak and the compatibility with the experimental data can be noticed in Fig. 26 and Fig. 27.



Fig. 31 Residual peak of the final spectrum



Fig. 32 Residual peak of the final spectrum

1.4 Spectral line identification

We tried to match the observed spectral peaks with the known energies of some elements (Fig. 28), which may actually be found in the materials of the setup (ceramics of the circuits, soldering, materials of the same setup,.....) as this also affords a rough check of detector linearity. We conclude that the detectors display a good linearity also for energies outside the interval 7000-10000 keV for which the first calibration was made.



CCD Spectrum current

Fig. 33 Possible elements

2. Charge transport correction

The experimental beta value is directly related to the resolution of the detector, so, in order to improve this value, an improvement of the detector resolution was carried out by including the *charge transport correction* in the experimental data. CCD is a device where an electrical charge is captured and converted into a digital value. This is achieved by "shifting" the signals pixel by pixel on the device, but during the transfer of the charge, part of it is lost (an illustration of the CCD charge transport is shown in Fig. 34).



Fig. 34 Concept of the charge transport in CCD [5]

The VIP CCD's are made of 1248×1152 pixels. The dependence of the ADC read value with the coordinate X and Y is plotted in Fig. 35 and Fig. 36.





Fig. 35 ADC dependence with x coordinate



Fig. 36 ADC dependence with y coordinate

For the correction algorithm, the profile of the two bidimensional histograms (Fig. 37 and 38), with a

cut of 3σ around Cu K_a line (obtained from the uncorrected data fit) are generated. From these profiles, using a linear fit (Fig. 39 and 40), the correction parameters are obtained.



Fig. 37 Profile cut with x coordinate



Fig. 38 Profile cut with y coordinate







Fig. 40 Profile y fit

The correction algorithm was applied for all set of data, therefore, each hard disk provided 16 set of parameters (slope and offset) for both, x and y coordinate. With these parameters the analysis previously described was repeated. Fig.41 shows the fitted histogram of the ccd 2, hard disk 3, where an improvement of the σ K_{α} of almost 30% can be noticed, by comparison with Fig. 11.



Fig. 41 CCD2, hard3 corrected channel spectrum

The overall improvements of the CCDs resolution (around 20%) are plotted in the Fig. 42 - 45. The charge correction algorithm did not alter the spectra, as result, using the same scale factor for the background normalization, the overlaying of the two histograms (Fig. 50) remains almost the same.



Fig. 42 Resolutions of K_{α} and K_{β} lines, for each hard disk signal data



Energy resolution nocurrentcor data

Fig. 43 Resolutions of K_{α} and K_{β} lines, for each hard disk background data



Fig. 44 Resolutions of K_{α} and K_{β} lines, for each CCD signal data



Energy resolution nocurrentcor data

Fig. 45 Resolutions of K_{α} and K_{β} lines, for each CCD background data



Fig. 47 Signal spectrum - zoom



Fig. 48 Background spectrum

Background spectra fit



Fig. 49 Background spectrum - zoom



Fig. 50 Corrected background and signal spectra

The charge transport correction gave a 20% improvement for the detector resolution translated also in an analogous $\beta^2/2$ betterment. Now the difference of events between the measurements with and without current, into the RoI, is $\Delta NX = 37\pm146$. Taking as a limit of observation three standard deviations, we get for the PEP violating parameter:

$$\beta^2/2 \le 3.25 \times 10^{-29}$$
 at 99.7 C.L. (14)



Fig. 51 Subtracted spectra full



Fig. 52 Subtracted spectra full - zoom



Fig. 53 Subtracted spectra full - beta

Final VIP subtracted spectra



Fig. 54 Subtracted spectra - zoom





Fig. 55 Elements

Although the energy calibration was done using the range 8-9 keV, the charge transport correction didn't move the peaks of the corresponding elements (Fig. 55).

After the charge transport correction the difference between K_{α} lines signal – background was slightly increased, increasing the visibility of the residual peak at 8 keV in the subtraction histogram, confirmed also by the differences of the fit function histograms in Fig. 56 and 57.



Fig 56 Subtracted fit functions histograms with errors from fit



Fig 57 Subtracted fit function histograms

Bibliography

- [1] Dirac-Fock method for muonic atoms: transition energies, wave functions, and charge densities, J.V. Mallow, A.J. Freeman and J.P. Desclaux, Phys. Rev A, 17, 1804 (1978).
- [2] Relativistic multiconfiguration Dirac-Fock package, *J.P. Desclaux*, http://dirac.spectro.jussieu.fr/mcdf/autres/24_mettec.html5
- [3] Ramberg, E., Snow, G.A.: Experimental Limit on a Small Violation of the Pauli Principle. Phys. Lett. B 238, 438 (1990).
- [4] Bartalucci, S., et al.: New experimental limit on the Pauli Exclusion Principle violation by electrons. Phys. Lett. B 641, 18 (2006)
- [5] http://www.microscopyu.com/articles/digitalimaging/ccdintro.html