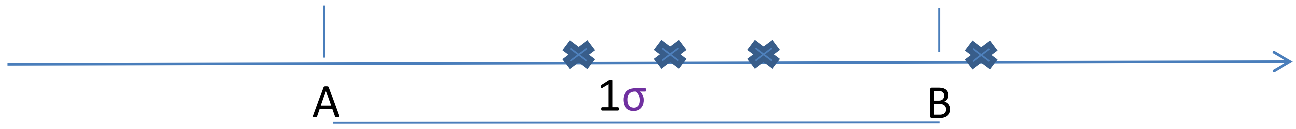
Particle identification method based on simultaneous checks of the variance and the average of the measurements of the Čerenkov angles of the photons emitted in the RICH radiatior(s). Its application in experiments performed in Hall A at Jlab

(G.M. Urciuoli)

General Concept



Let us suppose we have to measure a physical quantity which can assume only two values: “A” and “B” and we have at disposal a limited set of measurements. It could happen that the resolution “σ” of our experimental apparatus is not good enough to make us able to determine, simply calculating the average of the measurements, which is the value (“A” or “B”) the physical quantity had during our measurements.

In the example quoted above, the four measurements, marked with the sign “✖“, provide an average value closer to “B”, but we cannot exclude the possibility that the value the physical quantity had during our measurement was “A” instead, even considering that now, because we are dealing with the average of the measurements, the standard deviation has decreased from σ to σ/2. However, matters change drastically when we consider measurement distributions around the expected values “A” and “B” . In fact, it would have been very unlikely that all the single four measurements provided values bigger than “A”, if the physical quantity had this value during our measurements. In particular, a test on the variance “σ^2”, that is a test which make a comparison between the known variance of our apparatus, “σ^2”, and the sums $\sum(X_i - A)^2$ and $\sum(X_i - B)^2$, shows that the probability that the physical quantity had the value “B” is much bigger than it had the value “A”. This test on the variance and the test performed on the measurement average are completely independent and the two tests can hence be performed simultaneously, increasing the confidence level in the choice between “A” and “B”.

An example: check on a weight with a scale

Let us suppose we want to check if the weight of an item M is $X_{exp} = 10\ Kg$ by a weight scale. We know the scale resolution because we have performed a series of measurements on samples of known masses. We have assumed its resolution equal to the standard deviation σ_{exp} of these measurements, related to the variance σ_{exp}^2. The first and most straightforward test is to weigh N times the item M, calculate the average of these measurements:

$$\bar{X} = \frac{\sum_{i=1}^{N} X_i}{N}$$

and verify is $\bar{X}$ is statistically compatible with $10\ Kg$.

In addition, however, we can calculate the variance of the N measurements X_i:

$$\sigma^2 = \sum_{i=1}^{N} \frac{(X_i - \bar{X})^2}{N-1}$$

(When $N \to \infty$ $\lim_{N\to\infty} N-1 = N$; $\lim_{N\to\infty} \bar{X} = X_{exp}$)

We can estimate the variance through the sum:

$$\sigma_{est}^2 = \sum_{i=1}^{N} \frac{\left(X_i - X_{exp}\right)^2}{N\ -1}$$

$$\lim_{N\to\infty} \sigma_{est}^2 = \sigma_{exp}^2$$

If our assumption that the weight of our item M is $X_{exp} = 10\ Kg$ is correct, σ_{est}^2 must verify whatever statistical check, which compares it with σ_{exp}^2, the expected variance value.

The test on the variance and the test on the average of a set of measurements are completely independent

The sum:

$$\sum_{i=1}^{N} \frac{\left(X_i - X_{exp}\right)^2}{\sigma^2{}_{exp}} \qquad (1)$$

has N degrees of freedom, because no parameter has been deduced from the measurements X_i, as the value of X_{exp} has been supposed by us and $\sigma^2{}_{exp}$ has been deduced by previous measurements.

The sum can be transformed into:

$$\sum_{i=1}^{N} \frac{\left(X_i - X_{exp}\right)^2}{\sigma^2{}_{exp}} = \frac{1}{\sigma^2{}_{exp}} \times \left\{\sum X_i^2 - 2 \times \bar{X} \cdot X_{exp} + N \cdot X_{exp}^2 + N \cdot \bar{X}^2 + N \cdot \bar{X}^2 \ - 2 \cdot N \cdot \bar{X}^2\right\} =$$

$$\sum_{i=1}^{N} \frac{(X_i - \bar{X})^2}{\sigma^2{}_{exp}} + \frac{N}{\sigma^2{}_{exp}} \cdot \left(\bar{X} - X_{exp}\right)^2 \qquad (2)$$

In (2), the first term is a sum with $N-1$ degrees of freedom, as the measurement X_i are connected to each other by the value of $\bar{X}$. The first term is proportional to the variance. A test on it is hence a test on the variance.

The second term in (2) has only one degree of freedom and corresponds to the distribution of the average value $\bar{X}$ around the expected value X_{exp} with a standard deviation $\sigma_{exp}/\sqrt{N}$. A test on it is hence a test on the average value.

Because of the way they are derived, the two tests are independent (this can be verified by computer simulations too).

A test on the variance Is completely independent on a test on the average. As a consequence:

even if the hypothesis that M weighs $10\ Kg$ cannot be rejected considering the test on the measurement average only, the test on the variance could still prove this hypothesis is false. Even if the test on the variance cannot prove, on its own, that the hypothesis that M weighs $10\ Kg$ is false, we can ascertain the contrary considering simultaneously the test on the average and the test on the variance of the measurements. Checking simultaneously the average and the variance of a set of measurements helps hence when we can perform only a small number of measurements.

If, for example, we performed N measurements with our scale to weigh our item M, and we obtained the values $\bar{X}$ and σ^2 as the average and the variance of these measurements respectively:

a) If we obtain a probability smaller than 10% that $\bar{X}$ is statistically compatible with $10\ Kg$.
b) If we obtain a probability smaller than 10% that σ^2 is statistically compatible with our scale resolution.
c) The probability that the weight of M is equal to $10\ Kg$ is smaller than 0.1x0.1 < 1%.

In case of measurements following a Gaussian distribution, their variance can be easily checked with a χ^2test.

If the measurements of our weight scale follow a Gaussian distribution, thus each single measurement X_i performed by us, will follow a Gaussian distribution centred around the expected value X_{exp}. Then the sum:

$$Sum = \sum_{i=1}^{N} \frac{\left(X_i - X_{exp}\right)^2}{\sigma^2{}_{exp}} \qquad (3)$$

follows the χ^2 distribution with N degrees of freedom (no parameter derived by the N measurements).

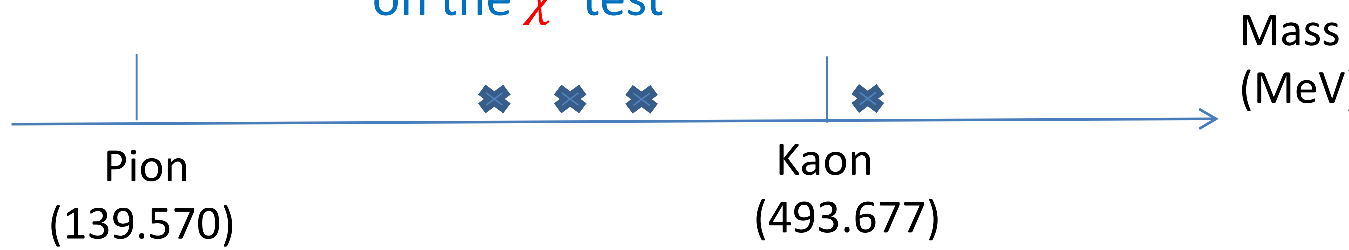
Checking the variance with a χ^2 test we can define a confidence level α by which we can accept or reject an hypothesis. For example setting our confidence level $\alpha = 0.001$, we can reject the hypothesis our item M weighs $X_{exp} = 10\ Kg$ if the probability:

$$\mathrm{P}(\chi^2 = Sum)_N < \alpha$$

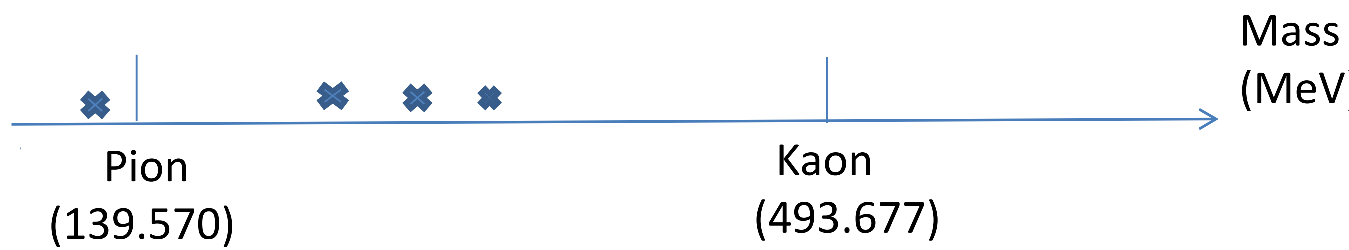
with $\mathrm{P}(\chi^2 = Sum)_N$ the probability that the χ^2 distribution with N degrees of freedom is bigger than the value of Sum provided by (3).

A Particle Identification based on the Maximum Likelihood can never perform better than a Particle Identification based on the χ^2test

Set of measurements "α"



Set of measurements "β"



Let us suppose a Kaon crossed the RICH and we detected 4 Čerenkov photons that provided a set of 4 values of its mass. The configuration labelled as Set of measurements "α" has a much bigger probability to happen than its mirror symmetric configuration labelled as Set of measurements "β". The χ^2test will correctly identify the particle which crossed the RICH, in case the configuration labelled as Set of measurements "α" occurs and will fail, vice versa, if the configuration labelled as Set of measurements "β" occurs. If a Maximum Likelihood Method correctly identifies the particle crossing the RICH as a Kaon, when the configuration labelled as Set of measurements "β" occurs, it will fail, for symmetry reasons, to identify the particle as a Kaon in case the configuration labelled as Set of measurements "α" occurs. Because the configuration labelled as Set of measurements "α" has a bigger probability to happen than the configuration labelled as Set of measurements "β", the Maximum Likelihood Method fails to identify correctly the particle more often than the χ^2test. This happens because the χ^2test is anchored to the event probability distribution, while the Maximum Likelihood is not.

The JLab Hall A RICH

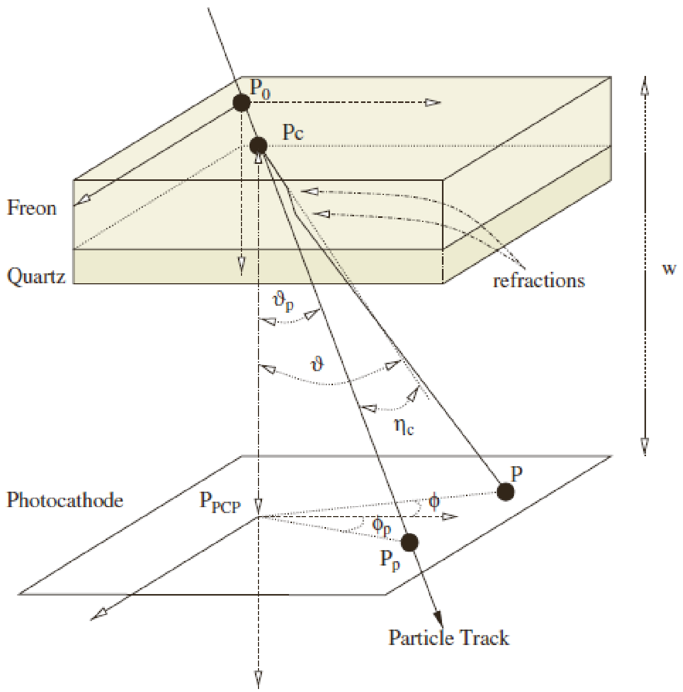


Fig. 5. Reference system used in the backtracking; (x_p, y_p) = MIP impact coordinates projected onto the pad plane; (θ_p, ϕ_p) = MIP polar and azimuthal angles; (x, y) = photon impact coordinates projected onto the pad plane; (θ, ϕ) = photon polar and azimuthal angles.

The RICH is similar to the experiment ALICE one. Čerenkov photons were generated in a 15 mm thick radiator made of liquid perfluorohexane (refractive index n = 1.29). After crossing a quartz window, Čerenkov photons entered the gap of a multiproportional chamber. This gap was filled with methane. At last, Čerenkov photons hit the chamber cathode made up by a number of cathode planes covered by 300 nm CsI film that acted as a photon converter. The generated electrons migrated towards the anode producing avalanches which produced clusters of fired cathode pads around the Čerenkov photons impact points on the cathode itself.

- From the barycentre coordinates of the clusters generated on the pad plane by the Čerenkov photons produced by a particle in the radiator, and from the particle direction and impact point coordinates on the multi-proportional cathode as determined by drift chambers, we determined the Čerenkov photon emission angle and hence the particle speed. From that, knowing the particle momentum as measured by a magnetic spectrometer we were able to determine the particle mass and hence to identify the particle itself.

An example: the Jlab Hall A RICH used in the experiment E94-108: (hypernuclear spectroscopy)

In the case of the experiment E94-107 in Hall A at Jlab, the detected particle momentum was 1.96 GeV/c and the refractive radiator index was n = 1.29. The Čerenkov photon angles for protons, K^+ and π^+ were respectively:

$$\vartheta^{p}_{exp} = 0.5366 \text{ rad};$$
$$\vartheta^{\pi^+}_{exp} = 0.6645 \text{ rad};$$
$$\vartheta^{K^+}_{exp} = 0.6807 \text{ rad};$$

The variances of the Čerenkov angle distributions around these three distributions were equal to:

$$(\sigma^{p}_{exp})^2 = (\sigma^{K^+}_{exp})^2 = (\sigma^{\pi^+}_{exp})^2 = \sigma^2{}_{exp} = 0.0174 \text{ rad}$$

To identify a particle which crossed the RICH and generated N clusters on the cathode of the proportional chamber, we calculated, for each cluster, the emission angle of the corresponding Čerenkov photon. We obtained in this way N measurements ϑ_i of the particle emitted Čerenkov photon angle. We calculated then the average value $\bar{\vartheta}$ of the N measurements ϑ_i:

$$\bar{\vartheta} = \frac{\sum_{i=1}^{N} \vartheta_i}{N}$$

And the three sums:

$$(\chi^p)^2 = \sum_{i=1}^{N} \frac{\left(X_i - \vartheta^{p}_{exp}\right)^2}{\sigma^2{}_{exp}}; \quad (\chi^{K^+})^2 = \sum_{i=1}^{N} \frac{\left(X_i - \vartheta^{K^+}_{exp}\right)^2}{\sigma^2{}_{exp}}; \quad (\chi^{\pi^+})^2 = \sum_{i=1}^{N} \frac{\left(X_i - \vartheta^{\pi^+}_{exp}\right)^2}{\sigma^2{}_{exp}};$$

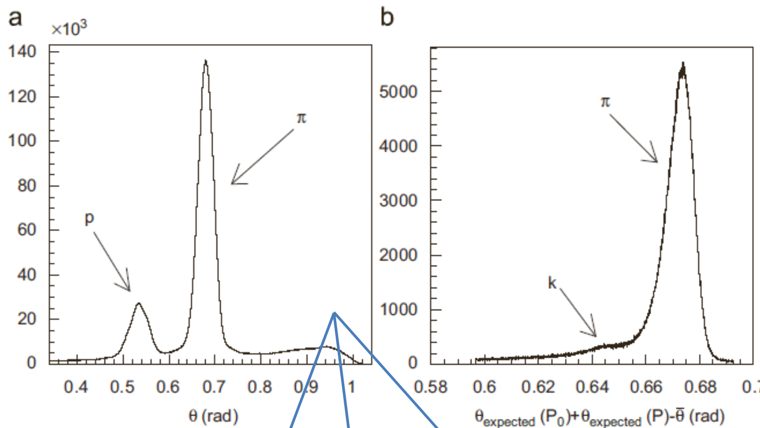
To identify a particle detected as K^+, we set a confidence level α_{rej} to consider not acceptable the values of $(\chi^p)^2$ and of $(\chi^{\pi^+})^2$, a confidence level α_{acc} to consider acceptable the value of $(\chi^{K^+})^2$ and a confidence level $\alpha_{accvalmed}$ to consider acceptable the average value $\bar{\vartheta}$ when checked against the expected Kaon Čerenkov emission angle. $\vartheta^{K^+}_{exp}$. Typical values were $\alpha_{rej} = 0.0001$, $\alpha_{acc} = 0.001$ and $\alpha_{accvalmed} = 0.001$.

The χ^2test as a method to eliminate false signals generated by noise

Very often we are in presence of "false" signals generated by noise/background. Most of these false signals can be identified and eliminated thanks to the χ^2test. In fact, let us suppose that none of the three sums: $(\chi^p)^2$, $\left(\chi^{K^+}\right)^2$ and $\left(\chi^{\pi^+}\right)^2$ is acceptable. Because, in any case, at least one particle has crossed the RICH, and hence one of the three χ^2 values would have been statistically acceptable, at least one of the signals we are analysing is false. One can hence eliminate 1, 2, …, $N_{removed}$ terms from the sums which define $(\chi^p)^2$, $\left(\chi^{K^+}\right)^2$ and $\left(\chi^{\pi^+}\right)^2$, starting from the biggest ones, until at least one of the three χ^2values is acceptable. Removing the biggest terms from the sums which define $(\chi^p)^2$, $\left(\chi^{K^+}\right)^2$ and $\left(\chi^{\pi^+}\right)^2$ enhances the probability to eliminate false signals instead of the true ones, because noise signals spread evenly in the RICH while true signals cluster around the expected value. In the hypernuclear spectroscopy experiments we were able to eliminate noise signals which amounted to 25% of the total signals. In the transversity experiment we were able to eliminate noise signals which amounted to 75% of the total signals. With method based on the calculation of the average of the signals or on the Maximum Likelihood, the elimination of noise signals is impossible.

Results (experiment E94-108):

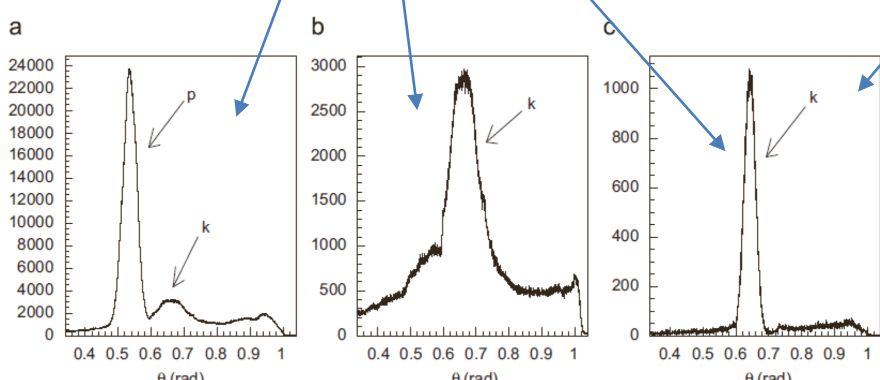
Inefficacy of the signal average calculation: in presence of noise:



ϑ_i distribution

$\bar{\theta}$distribution

χ^2test effectiveness (without average value test):



ϑ_i distribution

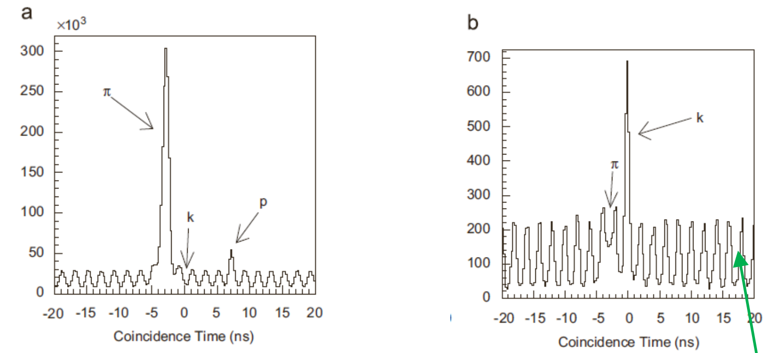
$\left(\chi^{\pi^+}\right)^2$ probability < 0.0001

ϑ_i distribution, both $\left(\chi^{\pi^+}\right)^2$and $(\chi^p)^2$ probabilities < 0.0001

ϑ_i distribution, both $\left(\chi^{\pi^+}\right)^2$and $(\chi^p)^2$ probabilities < 0.0001 and $\left(\chi^{K^+}\right)^2$ probability > 0.1

π^+rejection $\sim 100\%$;
K^+ detection efficiency $\sim 70\%$

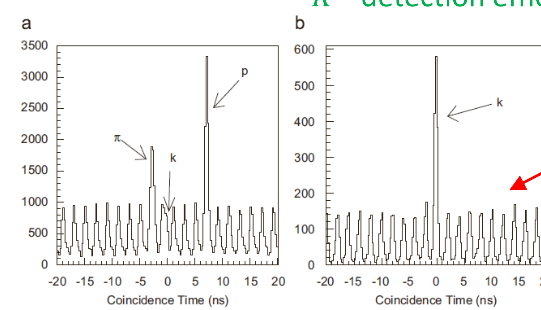
Effectiveness of the χ^2test and of the test on the average performed simultaneously:



No RICH PID

Both $\left(\chi^{\pi^+}\right)^2$and $\bar{\vartheta}-\vartheta^{\pi^+}_{exp}$ probabilities < 0.0001 and both $\left(\chi^{K^+}\right)^2$ and $\bar{\vartheta}-\vartheta^{K^+}_{exp}$ probabilities > 0.1

$\pi^+/_{K^+}$ rejection ratio $= 340 \pm 11$;
K^+ detection efficiency $\sim 91\%$



Two Čerenkov detector cuts only

Two Čerenkov detector cuts; $\left(\chi^{\pi^+}\right)^2$probability < 0.01; $(\chi^p)^2$ probability < 0.0001 and $\left(\chi^{K^+}\right)^2$ and both $\left(\chi^{K^+}\right)^2$ and $\bar{\vartheta}-\vartheta^{K^+}_{exp}$ probabilities > 0.001

π^+rejection $\sim 100\%$;
K^+ detection efficiency $\sim 96\%$

Conclusions:

- When dealing with quantized physical quantities which can assume only discrete values, it is not worthwhile to check the average of the measurements only. In fact, in principle, the average can assume any numeric value and hardly can match the expected discrete values. Besides, when performing average calculations, you lose the pieces of information provided by the single measurements.
- The variance of measurements instead, in principle can assume any value (it is not a quantized value). A check on its value is hence more performing than a check on the measurement average. Above all, a test on the variance and a test on the average are completely independent and can be performed hence simultaneously multiplying the rejection factors.
- A method based on the Maximum Likelihood can never match a method based on the variance. In fact, the latter, anchored to the statistical distribution, employs precise statistical tests while the former cannot do the same.
- A method based on a test on the variance allows us to eliminate the noise.
- When the measurements follow a Gaussian distribution around the expected value, a χ^2 test can be used to test the variance.
- All of above is valid whatever method (backtracing, cluster barycentre determination, etc.) is employed to determine the single photon Čerenkov angle emission.