A Kalman Filter approach for track reconstruction in a neutrino telescope

G. De Rosa *, Y. Pethukov**

*Univ. Federico II and INFN Naples, Italy

** *JINR,Dubna, Russia*

Track reconstruction in a Cherenkov neutrino telescope

- electromagnetic showers
- Cherenkov photon scattering
- ^{40}K beta-decay background

The arrival time depends on the track parameters (time, direction and position) as well as on the PMTs position (i.e. the detector geometry) and track reconstruction deals in a non-linear problem with VERY non-gaussian measument errors

We have non-linear problem with non-gaussian distributions

Kalman Filter technique

uses measurements observed over time that contain noise and other inaccuracies

In contrast to other methods, the process parameters evaluation proceeds progressively including the information of each additional measurement, thus improving iteratively the knowledge on the current parameters of process.

- Updated for each new measurement,
- Needs informations from previous state only, not all previous measurements, so, used measurements can be discarded,
- Saves computation capacity and storage,
- Proper (and relatively simple) dealing with noises,
- Possible to propagate information from last to first measurement (smoothing)

Kalman Filter - state

The track is regarded as a dynamic system described by the **state vector** x_k for k-th hit position

The evolution of the state vector is described by a **System equation**

$$x_k = f_k(x_{k-1}) + \omega_k \, ,$$

where ω_k is noise due to multiple scattering and energy losses with covariation matrix Q_k ; f_k is the track propagator from hit $k-1$ to hit k (track model);

With $F_k = \partial f_k / \partial x_k$, first order gives us (**linear approximation**)

$$x_k = F_k x_{k-1} + \omega_k$$

Kalman Filter - measurement

The **measured state vector** is obtained from the quantities measured by the k_{th} hit, m_k, that are functions of the state vector,

$$m_k = h_k(x_k) + \varepsilon_k \quad ,$$

corrupted by a **measurement error** ε_k with covariation matrix V_k ;

With $H_k = \partial h_k / \partial x_k$, first order gives us (**linear approximation**, Extended Kalman Filter, EKF)

$$m_k = H_k x_k + \varepsilon_k$$

Kalman Filter
prediction and filtering (add information)

State vector prediction in «k-th» position is based on state vector in «k-1» position

$$x_{k|k-1} = f_k(x_{k-1}) + \omega_k \quad ,$$

and prediction covariation matrix

$$C_{k|k-1} = F_k C_{k-1} F_k^T + Q_k$$

The filter step updates the state vector on the basis of the all previous measurements

$$x_k = x_{k|k} = x_{k|k-1} + K_k(m_k - H_k x_{k|k-1}) \ ,$$

where Kalman gain matrix

$$K_k = C_{k|k-1} H_k^T (V_k + H_k C_{k|k-1} H_k^T)^{-1}$$

and filtered covariation matrix in «k» position

$$C_k = C_{k|k} = (I - K_k H_k) C_{k|k-1}$$

The track model

The evolution of the state vector depend on the track parameters via the track model f_k.
In the case under study, we set:

$$f_k(\bar{x}_{k-1}) = \bar{x}_{k-1} + w_k$$

Notice that, for energetic muons, multiple scattering has a negligible effect; the energy loss is not taken into account at this step, thus the process noise w_k **is neglected**

the track is considered as a straight line and is the same for all hits, not depending on hit position

The track model

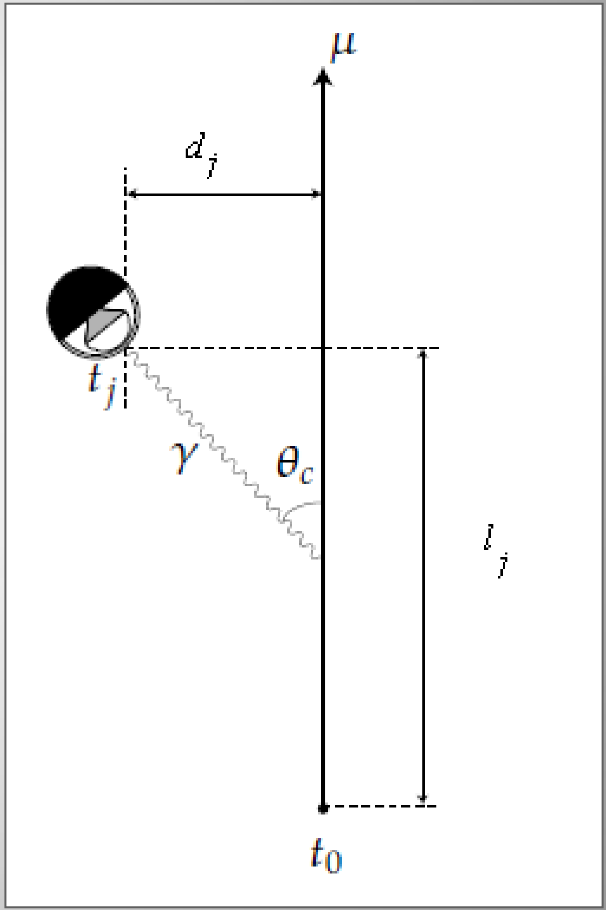
In the Extended Kalman Filter approach, the measurement function h_k can be parametrized as:

$$h_k(\bar{x}_k) = T_{theor}$$

using the relation

$$(t_j - t_0) = \frac{1}{c}\left(l_j - \frac{d_j}{\tan\theta_C}\right) + \frac{1}{v_{ph}}\frac{d_j}{\sin\theta_C}$$

for the evaluation of T_{theor}



Kalman filtering with shifted mean (KFS)

Since the Extended Kalman filter strategy requires linear approximation of the track model only over a short range, it is suitable for measurements in a neutrino.

Photons originating from the muon and arriving at the OM without scattering in the water carry the most precise timing information: their arrival time is only perturbed by dispersion and the TTS of the PMT and hence the residual distribution is sharply peaked at 0.
Photons that originate from secondary electrons or that have scattered, are in general delayed with respect to this time.

Kalman filtering with shifted mean (KFS)

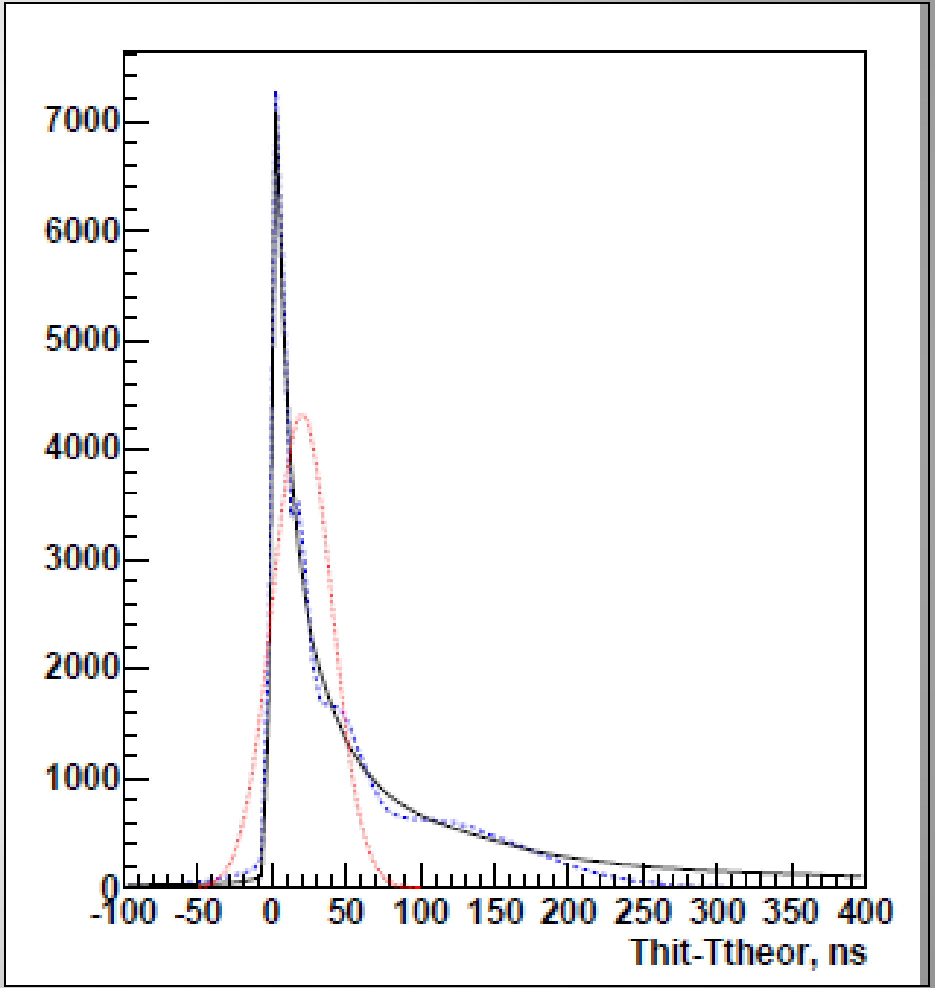
we describe residuals distribution with a Gaussian, but with mean shifted to some appropriate value r > 0

Measurement function:

$$h_k(\bar{x}_k) = T_{theor} + T_{shift}$$

T_{shift} is a time shift (we obtain optimal performance at 20 ns)

NOT good results obtained with this approach!



solid — Monte-Carlo,
red dot — «best» approximation with one gaussian
blue dash — approximation with sum of five gaussians,

Limitation of KF

The problem we are dealing is not Gaussian!

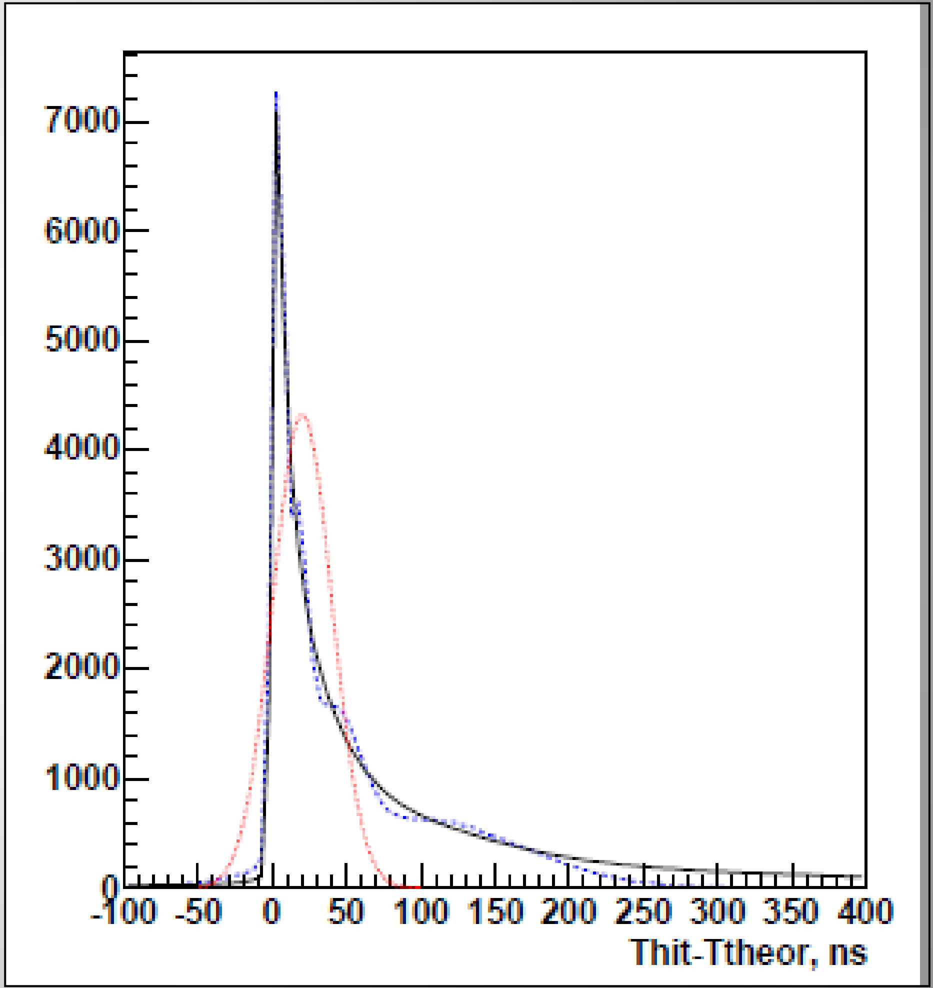
Actually, KF is not the best method when we have non-Gaussian distributions:

- hit errors might have tails due to noise, overlapping events, etc.
- multiple scattering has larger tails than Gaussian
- ionization energy loss follows Landau distribution
- electron and high energy muon energy loss (bremstrahlung) follows Bethe-Heitler distribution

Gaussian Sum Filter (GSF)

A method that takes better into account non-Gaussian distributions of measurement errors is the Gaussian-Sum Filter

The distribution of the measurement errors may be approximated by sum of several Gaussian distributions, so that a Gaussian mixture model becomes the natural description: the core corresponds to the principal component of the Gaussian mixture, and the tails can be modeled by one or several additional Gaussians.



solid — Monte-Carlo,
red dot — «best» approximation with one gaussian
blue dash — approximation with sum of five gaussians,

Gaussian Sum Filter (GSF)

so, approximating non-gaussian measurement errors distribution as sum of several gaussians:

$$p(\epsilon_k)=\sum_1^m p_i \varphi(\epsilon_k,\mu_i,\sigma_i) \quad , \quad \sum_1^m p_i=1$$

State vector becomes multicomponent, because one measurement produce one estimation of state vector for each measurement error gaussians

$$x_{k|k}^{ij}=x_{k|k-1}^{j}+C_{k|k}^{ij} H^T \sigma_i^{-2}\left(y_k-H\,x_{k|k-1}^{j}-\mu_i\right)$$

with covariation matrix

$$C_{k|k}^{ij}=\left(\left(C_{k|k-1}^{j}\right)^{-1}+H^T \sigma_i^{-2} H\right)^{-1}$$

Gaussian Sum Filter

State vector now is a combination of state vector components

$$x_{k|k} = \sum_{l=1}^{n_k} q_k^l x_{k|k}^l$$

With corresponding covariation matrix

$$C_{k|k} = \sum_{l=1}^{n_k} q_k^l C_{k|k}^l + \sum_{l=1}^{n_k} \sum_{m>l}^{n_k} q_k^l q_k^m \left(x_{k|k}^l - x_{k|k}^m\right)\left(x_{k|k}^l - x_{k|k}^m\right)^T$$

and Weights

$$q_k^l = q_k^{ij} = p_i q_{k-1} \varphi\left(y_k; H x_{k|k-1}^j, \sigma_i^2 + H C_{k|k-1}^j H^T\right)$$

Index «l» (and «m») join indexes «i» and «j» in one sequence.

full result is weighted mean of accumulated components

Gaussian Sum Filter

we assume that the distribution of the measurement errors ε_k can be modelled by a Gaussian mixture with five components:

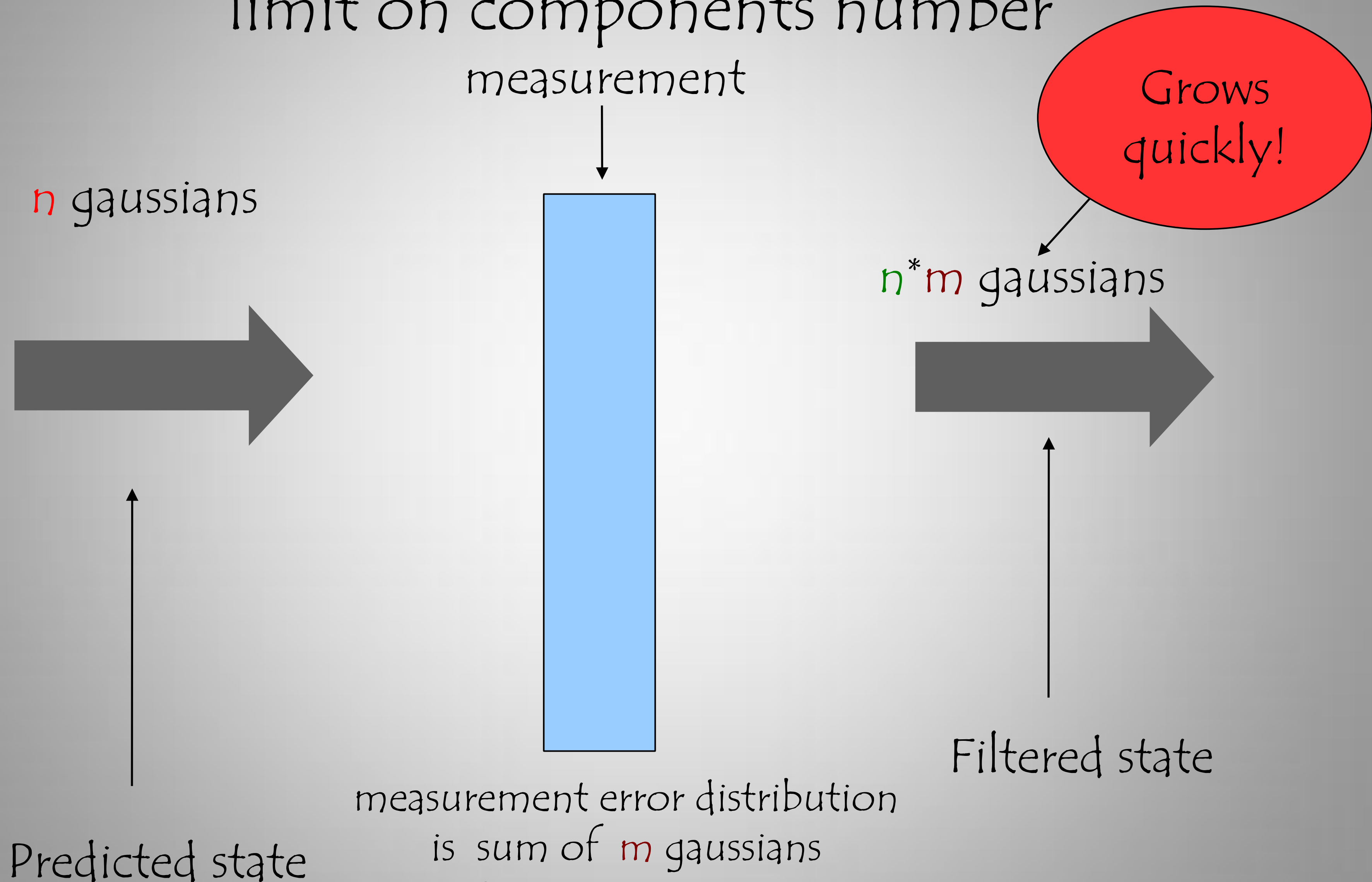
$$p(\varepsilon_k) = \sum_{i=1}^{5} p_i \cdot \varphi(\varepsilon_k; \mu_1, \sigma_i)$$

$$\sum_{i=1}^{5} p_i = 1$$

Distribution of measurement errors was obtained from Monte-Carlo data as distribution of measured hit arrival time and theoretical one ($m_k - T_{theor}$)

The GSF resembles a set of Kalman filters running in parallel, each Kalman filter corresponding to one of the components of the state vector mixture.
Each of the filters or components has a weight attached.

Gaussian Sum Filter: limit on components number



Gaussian Sum Filter limit on number of components

A strict application of the GSF algorithm quickly leads to a prohibitively large number of components due to combinatorics involved each time a hit is added

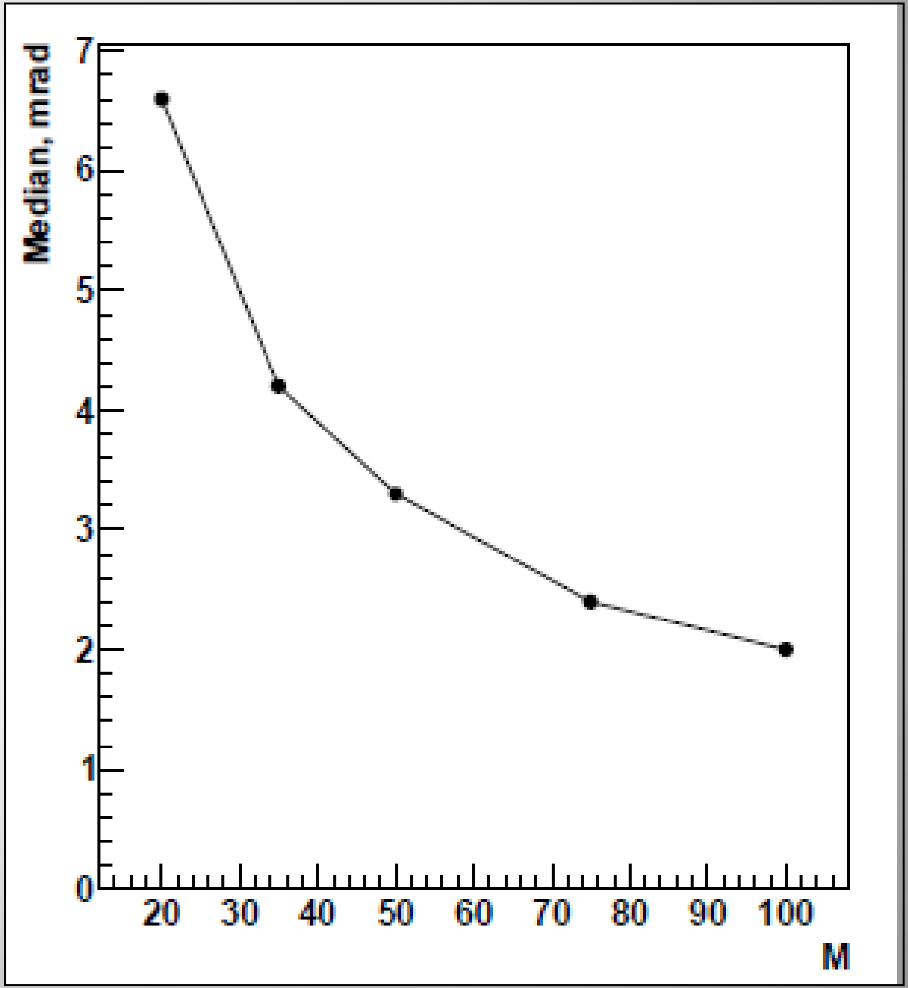
In each filtering step, we obtain 5 times more state vector components than the number of predicted ones: after filtering step for the k – th hit the number of components is 5^k.

Possible strategies:

- select M components with maximal weights,
- clusterization (on base of distance between components) ,
- resampling — random choice of M components with correspondance to them weights.

Gaussian Sum Filter
limit on number of components

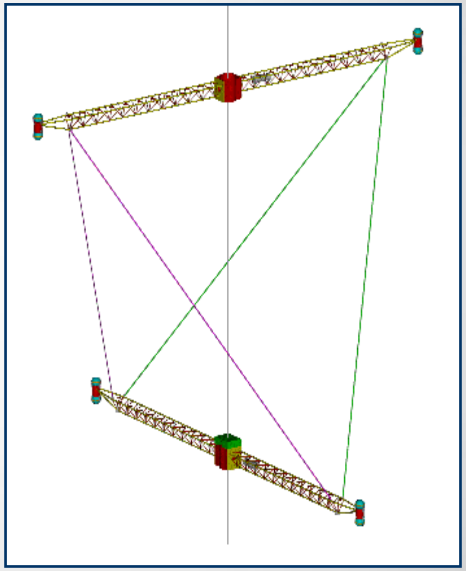
By denoting with M the maximum number of components allowed in the filtered step, we studied the dependence of the algorithm track angle error on M:
a value about of at least 50 components is quite good from point of view of the algorithm performances.

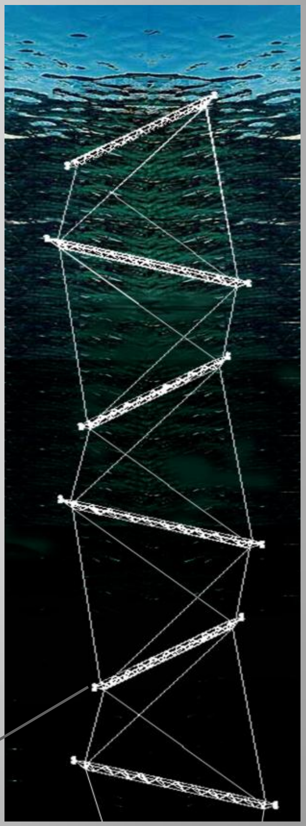


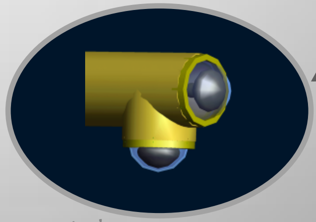
Results

We will refer, for this study, to the one proposed by the NEMO Collaboration as reference detector

A square array of structures (towers) composed of a sequence of "storeys" hosting the PMTs. Each storey will be rotated by 90°, with respect to the upper and lower adjacent ones, around the vertical axis of the tower:
-9×9 towers spaced 140 m, with 72 PMTs for each tower; 5832 PMTs in total
- detector volume ~ 0.9 km^3
- 18 storey per tower
- storey 20 m long with two optical modules (one downlooking and one looking horizontally) at each end (4 OMs per storey)
- Distance between storeys is 40 m

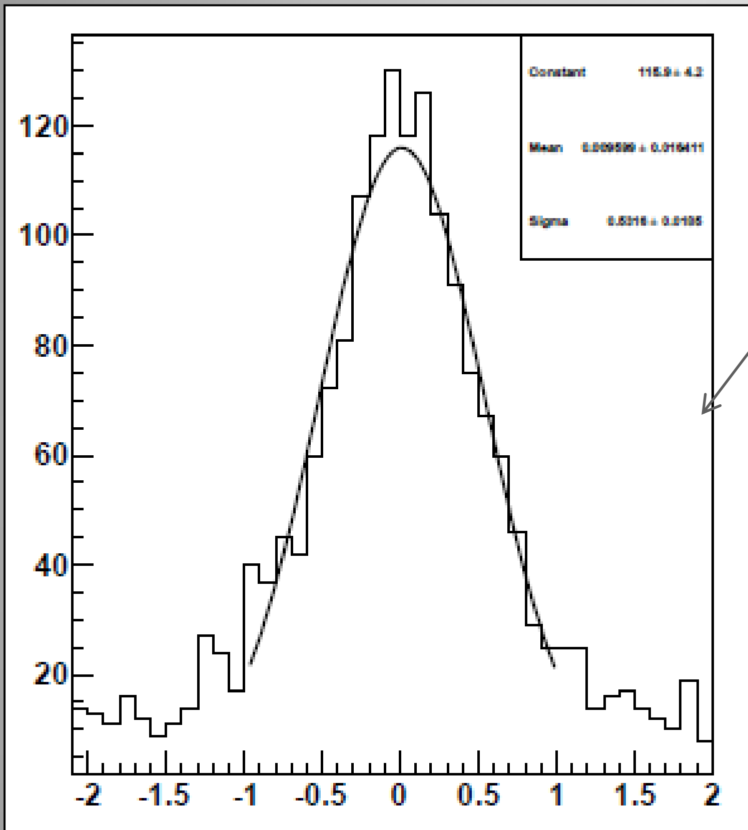




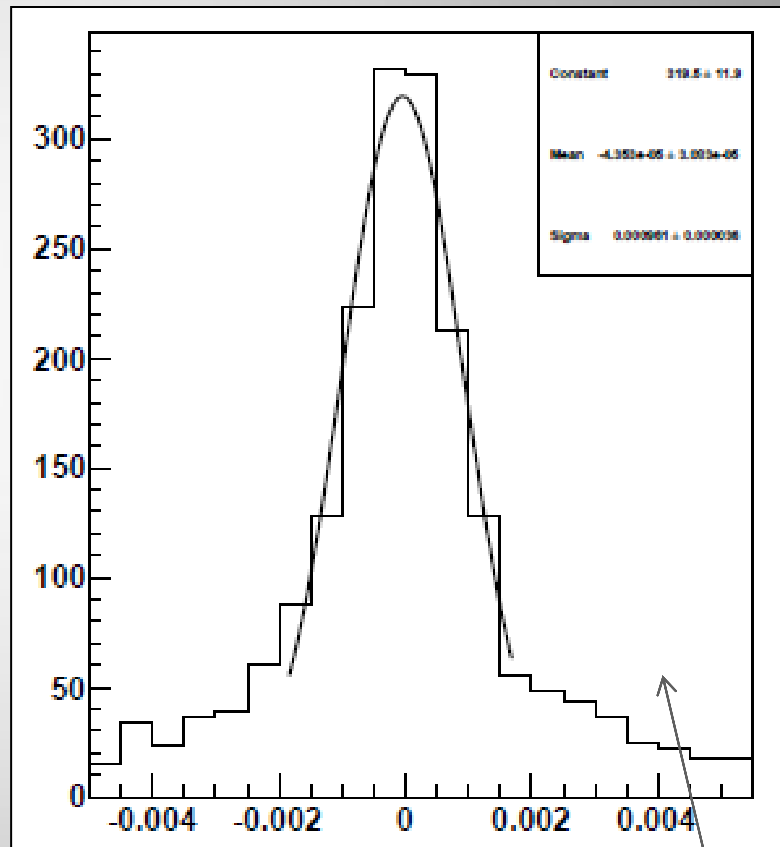


The muon flux is distributed as E^{-1} whitin a range 100 GeV-10^4 TeV. Muons are going upward and distributed uniformly in the hemisphere. Start points of the tracks are always placed outside of the detector.

Results



distribution of the difference between reconstructed and generated X-coordinate parameter



distribution of the difference between reconstructed and generated zenith angle parameter

Method	m_θ	σ_θ	σ_X
ANTARES	1.95	0.93	0.66
GSF	2.00	0.96	0.53
KFS	12.3	6.90	4.08

Results

- it works and gives parameters accuracies in agreement with the maximal likelihood ones,
- it is faster then maximal likelihood,
- improvements are possible (optimization),

The mean CPU speed per event (on Intel Xeon 2.66 Ghz processors) for the GSF algorithms is about 4 time faster than the maximum likelihood method.

Outlook

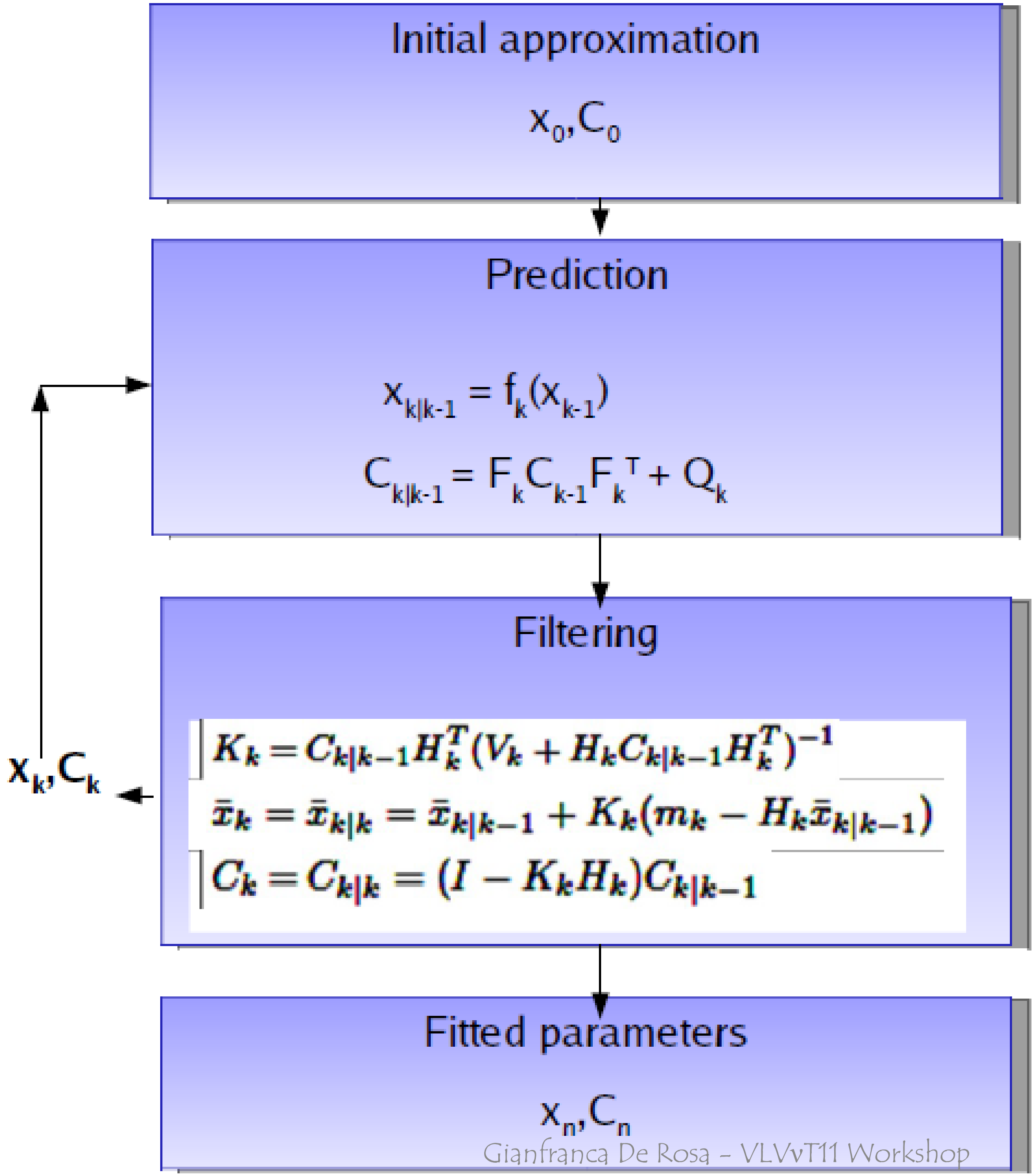
- Results on track reconstruction strategy based on Gaussian Sum Filter has been presented
- Effective area and angular resolution for a km3 neutrino telescope with this strategy are under study

- In addition, with this approach, also **energy loss** can be taken into account:
with the GSF reconstruction of muon tracks is possible to obtain a good Gaussian-mixture approximation of the bremsstrahlung energy loss

WORK in PROGRESS!

more than.....

Kalman filter



Process Noise Q_k

Measurement:
m_k, H_k, V_k

Kalman filter flowchart

Time update equations (predict)

$$x_{k|k-1} = f_k(x_{k-1})$$

$$C_{k|k-1} = F_k C_{k-1} F_k^T + Q_k$$

Predict state and covariance estimates forward in time

Measurement update equations (correct)

$$K_k = C_{k|k-1} H_k^T (V_k + H_k C_{k|k-1} H_k^T)^{-1}$$

$$\bar{x}_k = \bar{x}_{k|k} = \bar{x}_{k|k-1} + K_k(m_k - H_k \bar{x}_{k|k-1})$$

$$C_k = C_{k|k} = (I - K_k H_k) C_{k|k-1}$$

Compute Kalman gain K

Measure process m_k

Compute a posteriori estimate x_k

Compute a posteriori error covariance estimate C_k

Initial estimates

The track model

the parameters are the same for different hits, i.e. straight line is the same for all hits not depending on hit position.

In this case, the predicted error matrix

$$C_{k|k-1}$$

is equal to the filtered error matrix

$$C_{k-1}$$

for $\bar{x}_{k-1}$

since the matrix of derivatives Fk in the prediction covariation matrix

$$C_{k|k-1} = F_k C_{k-1} F_k^T + Q_k$$

is a unit matrix in our case

Gaussian Sum Filter limit on number of components

As starting point, we consider the state vector $\bar{x}_0$ and the covariance matrix C_0 which is an "infinite" covariance matrix, i.e. a large multiple of the identity matrix.
In each filtering step, we obtain 5 times more state vector components than the number of predicted ones, because we have different filtering values for different components of measurement errors distribution $p(\varepsilon_k)$.
The number of state vector components changes from 1 to 5^{k-1} for prediction step for $k-$ th hit. Thus, after filtering step for the $k-$ th hit the number of components is 5^k.

The initial track hypothesis

The initial track hypothesis, namely an initial value of the state vector together with its covariance matrix is derived from ANTARES prefit strategy

In the track fitting procedure, as preliminary step, the measured hit compatible with a signal are selected. Since hits with the largest charge are most likely related to the signal, each hit is chosen on the basis of its amplitude.

Also the timing and spatial distance with respect to the track hypothesis is taken into account. To discard background hits, the filtered 2 contribution of the hit is used as a criterion.

Bremsstrahlung energy loss

The optimal treatment of radiative energy loss within the Kalman filter formalism is to correct the momentum part of the state vector with the mean value of energy loss and to add the variance of the energy loss distribution to the relevant part of the covariance matrix

Approximating the Bethe-Heitler distribution with a single Gaussian is a quite crude approximation

The basic idea is to model the Bethe-Heitler distribution as a Gaussian mixture g(z) instead of a single Gaussian, the different components modelling different degrees of hardness of the bremsstrahlung in the layer under consideration

Bremsstrahlung energy loss

According to the description of the bremsstrahlung process, the particle retains its direction of propagation and a fraction z of its energy, with the probability density given by the Bethe-Heitler distribution.

This probability density $f(z)$ depends on the amount of traversed material and is highly asymmetric, with a singularity at z=1 and a long tail extending to very small z.