Chapter 1

Polfit

Polfit (**P**attern **o**f **L**ight **Fit**ter) is a newly developed track fitter. It works by comparing observed charge distribution with light pattern that would be generated for a given trial configuration of tracks. It varies track parameters (directions and momenta) until best match is found. It is different from other track fitters. It does not look for ring edges, instead it fits overall light pattern.

At the heart of the fitter is the likelihood "engine". It acts like a black box, taking as input hypothetical track configuration, and returning a number which tells us how closely the hypothesis matches what is observed. On top of the likelihood engine runs a fitting algorithm, which picks initial tracks and decides where next best candidate hypothesis lies, and finally decides when to stop fitting.

The likelihood engine consists of two major parts. The first part generates expected light distribution, given a track configuration. The second part compares the expected and observed light distributions and produces a number which is a function of how similar the two distributions are.

1.1 Expected Light

Expected light is generated in three stages. First direct light is generated, without any absorption or scattering. Then scattered light is added. Finally corrections for attenuation and geometry are added.

1.1.1 Direct Light

Direct light is generated for each track separately.

Cherenkov photon flux is assumed to be axially symmetric around track direction (no ϕ dependence). Azimuthal dependence is tabulated as a function of angle θ, with respect to track direction.

For very short tacks viewed from very far away this parameterization would be sufficient. This would be an approximation of zero-length point-like tracks viewed from infinite distance. However in Super-Kamiokande detector the length of the track (or electromagnetic shower) and nearness of the PMTs to the track cannot be ignored. Therefore the Cherenkov photon flux table depends on another variable r — the distance of the PMT from the vertex.

Finally the shape of the electromagnetic shower and its length depend on the momentum of the the initial particle — p. This is the third variable of the photon flux table.

There is a separate table for each particle type.

The Cherenkov photon flux tables were generated by Monte Carlo simulation. For each momentum value in the table many simulations of mono-energetic particles in water were run. Particle interactions (e.g. development of electromagnetic showers, scattering) were modeled. Charged particles generated Cherenkov photons. Cherenkov light did not scatter or attenuate. Cherenkov photon flux was measured at various distances and angles. For each distance and angle average flux from many simulations was calculated and saved in the table.

For electrons and gammas the flux table has 50 equally spaced values of momentum between 50 and 2000 MeV. There are 30 equally spaced values of r from 100 to 3000 cm. There are 200 equal size $\cos(\theta)$ bins from -1 to 1.

Photon flux is interpolated (extrapolated) linearly between (outside of) the values of p and r and centers of $\cos(\theta)$ bins.

1.1.2 Scattered Light

In the process of the development of the fitter it turned out that direct light is not enough. There are cases when direct light alone does not adequately describe observed light distribution. This discrepancy causes incorrect fit results.

Precise calculation of scattered light is very slow. However scattered light is a relatively small correction to the overall distribution of the light. Time can be saved by calculating it approximately.

It turned out that existing code for approximate calculation of scattered light was too slow. It required full calculation at every fitter step.

A new algorithm was developed which requires to perform most of the calculation only once, and only a simple fast calculation is performed at every fitter step, provided that the vertex does not move.

The most important simplifying assumption in this approach is that all direct light originates at the vertex.

A light ray emitted from the vertex in direction of any given PMT produces a specific pattern of scattered light on all other PMTs. If we change the intensity of the direct light beam the intensity of the scattered light pattern of course changes linearly, but the overall pattern stays the same. If we knew what scattered light pattern is produced for direct light emitted in direction of any

PMT we could simply add the patterns weighting them with how much direct light is expected to be seen by the each of the PMTs.

In other words there would be a matrix which tells us how much scattered light j-th PMT would see for direct light emitted towards i-th PMT. To get scattered light distribution we would multiply vector of direct light distribution by the scattered light matrix.

However such matrix would have $11146^2 = 124233316$ (over hundred million) entries. It would require large amount of computer memory and manipulating it would be slow.

Fortunately it turns out we don't need to consider directions towards every PMT. Scattered light pattern is smooth enough that we can consider much fewer directions.

The detector is divided into patches. For each patch the direct light seen by all PMTs which belong to the patch is added up. This makes up the direct light vector. After matrix multiplication we get scattered light vector. The amount of scattered light we get for each patch is assigned equally to each PMT in the patch.

The patches used for direct light and those used for scattered light don't have to be the same. In fact we use fewer direct light patches than scattered light patches. The scattering matrix is not square in such case. Relatively large direct light patches can be used when scattering is assumed to be isotropic.

Scattering matrix is calculated once before fitting iterations.

A light ray is considered leading from the vertex to the central PMT of a direct light patch. This light ray is a linear source of secondary light i.e. scattered light. The linear source gets dimmer exponentially further away from the vertex due to attenuation.

The source is viewed from a central PMT of a scattered light patch. To calculate amount of scattered light from the light ray reaching the PMT integration is performed along the linear source. The integrating algorithm samples amount of scattered light emitted towards the PMT from several discrete points along the linear source. The linear source is split into segments (one for each point) and each point represents total amount of scattered light from its segment. By default Romberg integration (reference - numerical recipes?) is used. In such case the line is divided into 9 equal segments with a point in the middle of each segment. A faster, but less accurate 3 point integration is available too.

At each point amount of scattered light is total amount of light attenuated in the corresponding line segment.

The total amount of light to be scattered is divided into two parts: isotropic part and angle-dependent part. The angle-dependent part represents Rayleigh scattering and scatters with angular distribution of $\cos^2(\theta)$, i.e. it is a forward-back scattering. By default angle-dependent fraction is set to zero i.e. only isotropic scattering is calculated.

The intensity of the scattered light seen from any point falls off as $1/r^2$. However this intensity becomes very large for small distances. The point represents a diffuse light source which occupies volume of a segment (1/9-th) of the solid angle of the direct light patch as seen from the vertex. To avoid the unwanted divergence r is restricted to values greater than $r_{\min}$. By default $r_{\min}$ = 400 cm, which is comparable with the size of direct light patch.

The primary light beam from the vertex to the central PMT of the direct light patch represents photon flux from the vertex to the whole patch. The flux is proportional to the solid angle of the patch as viewed from the vertex. Therefore the intensity of the light beam is adjusted taking into account distance to the patch and angle of the patch plane with respect to the light beam.

1.1.3 Corrections

Scattered light was calculated based on direct light photon flux. After that detector effects are added to the direct light. A series of corrections is applied which translate photon flux into charge observed by PMTs.

The largest correction is due to solid angle effect. PMT close to the vertex sees much more Cherenkov photons than PMT far away because it covers larger solid angle as seen from the vertex. Direct light is multiplied by factor $1 - r/\sqrt{r^2 + r_{\text{PMT}}^2}$, where r is distance of the PMT from the vertex, and r_{PMT} is PMT radius.

Another correction is for PMT acceptance variation with incident angle of the light.

Final correction is for water attenuation. Different wavelengths attenuate differently in water, but a single exponential describes effective attenuation well enough. Attenuation length used for the calculation is the measured effective attenuation length in Super-Kamiokande at the time when the fitted data was taken. In the case of Monte Carlo the same attenuation length as the one used in Monte Carlo simulation is used. Expected charge is multiplied by factor $e^{-r/\lambda}$, where λ is the attenuation length.

After the corrections are performed scattered light is added to the direct light.

1.1.4 Normalization

Final step of expected light calculation is normalization. Expected light is normalized by observed light.

Only observed PMT hits within specific timing window are counted (see section 1.3.2). Remaining PMTs are treated like they were not hit by any photon. By default PMTs with observed charge below 1 p.e. are assigned single p.e. value. This is because PMT which recorded a hit observed at least one photon. Bad PMTs are skipped (ignored) in both expected charge and observed charge arrays. After those adjustments all observed PMT charge is added up. Expected charge is multiplied by a factor which makes total expected charge identical to total observed charge.

Real PMT saturate at high charge. This has to be accounted for consistently in normalization calculation and later in likelihood calculation. A maximum charge value is used, by default 200 p.e., which is somewhat below typical saturation levels. Both expected and observed charge sums are recalculated again but with charges above the maximum charge counted as the maximum charge only. The expected charge is renormalized again to the "saturated" observed charge using the truncated sums. The procedure is repeated until the normalization converges.

1.2 Likelihood

At this point we have a normalized expected charge distribution. We can compare it to the observed charge distribution. In the comparison bad PMTs are ignored (skipped). Observed charge has a timing window cut applied (see section 1.3.2).

One of two comparison methods is used.

First method calculates for each PMT likelihood that for the given amount of expected charge the PMT records the observed amount of charge. Empirical formulas and lookup tables are used to calculate the likelihood value. Logarithm of the likelihood is taken and the log-likelihoods of all PMTs are added up (an operation equivalent to multiplying all likelihoods). Every individual likelihood is constrained to be above a minimum value, by default the minimum is 10^{-7}. The sum is the overall log-likelihood that tells us how well the observed and expected light distributions match. The number is always negative. The larger the number (less negative), the better is the match.

The second comparison method is the one used by default. A simple Poisson χ^2 formula is used (ref. PDG): $\chi^2_{\mathrm{PMT}} = 2(q_e - q_o) + 2q_o \ln(q_o/q_e)$. This simple formula turns out to give as good fit results as the likelihood method. By default before the calculation observed charge below 1 p.e. is set to 1 p.e. level for consistency with normalization calculation. Also by default when both observed and expected charges are above the saturation level used in normalization calculation (see section 1.1.4), the expected charge is set to the saturation level. Values of χ^2_{PMT} for every PMT are added. The total sum of the overall χ^2 that tells us how well the observed and expected light distributions match. The number is always positive. The smaller the number, the better is the match.

1.3 Fitter

Using the general likelihood engine described above fitters can be developed which search for configuration of rings which best matches observed pattern of light.

A specialized π^0 fitter was developed. It is designed to run on sample of events which were reconstructed as single showering ring. It tries to reconstruct two rings, each created by a gamma ray from the decay $\pi^0 \to \gamma\gamma$.

As input it takes vertex, total energy of both gamma rays and direction of the first gamma ray. Typically the input parameters are taken from the reconstructed vertex, visible energy and direction of the reconstructed ring, respectively. It fits the direction of the second gamma ray and fraction of the total energy taken by the second gamma.

1.3.1 Initial Guess

Before fitting starts an estimate is made where the plane is located in which the two gamma rays lie.

First timing and clustering cuts are performed on observed charge. Observed charge is then corrected for attenuation and PMT acceptance. The charge is projected onto a sphere centered around vertex. The projected charge is treated as a mass distribution and inertia tensor is calculated. The tensor is diagonalized and eigenvectors with corresponding eigenvalues are obtained.

The eigenvector corresponding to the smallest eigenvalue determines the direction along which most of the charge is located. The eigenvector corresponding to the largest eigenvalue determines the direction perpendicular to the plane in which most of the charge is located. In an ideal case with perfect Cherenkov rings projected on a sphere this plane would be the $\gamma - \gamma$ plane. Detector effects and fluctuations cause deviations from the ideal case.

1.3.2 Iterations

The fitter varies three variables: two angles ϕ and θ which parameterize direction of the second ring in spherical coordinate system, and fraction of total energy carried by the second gamma f. The spherical system is defined in such way that the first gamma lies on the equator with both ϕ and θ equal to zero. The poles at $\theta = \pm 90^{\mathrm{o}}$, the first ring, and the eigenvector corresponding to the largest eigenvalue lie in one plane. In such way the equator plane approximates the $\gamma - \gamma$ plane determined by the inertia tensor, given the constraint that the equator contains the first ring. Negative ϕ direction is chosen on the same side as the eigenvector corresponding to the smallest eigenvalue. This is the side of the first ring on which majority of charge is located and where we expect to find the second ring.

Before fitting proceeds timing cut is performed on observed charge PMT hits. First time of flight of light from the vertex to every PMT is subtracted. Sliding timing window of a constant width is moved until maximum number of hits is found. The width of the window is 30 ns for events with total number of p.e.'s corresponding to energy below about 1 GeV and it is expanded linearly for events with more charge.

Fitting is done using simplex method (ref. Numerical Recipes). For N-parameter fit in N-dimensional parameter space at every fitting step a set of $N+1$ points (called simplex) is considered. The simplex points cannot lie on a $N-1$-dimensional hyper-plane, i.e. (hyper)volume contained in a figure with corners defined by the simplex points must be non-zero.

The initial simplex (set of points) must fulfill this condition too. The simplest way to ensure this is to pick one point, and the remaining N points are set by moving the first point along one coordinate, different for each point.

By default the first point is picked at the coordinates $(\phi, \theta, f) = (-30, 0, 0.4)$ (angles are measured in degrees). The remaining points are chosen by changing one coordinate to one of the values $(20, 20, 0.2)$.

Fitting stops when the distance between the best and the worst simplex points is less than 1, where distance is measured using units for each axis defined by metric $(2, 2, 0.02)$.

1.3.3 Final Decision

In the χ^2 distribution for some events there is a narrow minimum in the direction of the first ring, i.e. when the two rings are almost on top of each other. This phenomenon occurs more strongly for electron events than for π^0 events.

In order to improve discrimination between electrons and pions a second fit is performed, following identical procedure. The only differences are that the points of the starting simplex are chosen at the vicinity of the first ring, and the stopping condition requires somewhat higher precision of the result. The first point of the simplex is $(0, 0, 0.2)$, the remaining simplex corners are chosen from the values $(5, 5, 0.3)$, and the metric for the stop condition is $(1, 1, 0.01)$.

This fit is called a "forward" fit, because it concentrates in the forward direction defined by the first ring. The first fit is called a "wide" fit because the initial simplex spans quite large area of the parameter space.

Finally a decision is made which fit is the best result. 100 is added to χ^2 of the forward fit and the fit with smaller χ^2 is chosen.