

Image optimisation techniques in a PET diamond location system

Martin Cook, Sergio Ballestrero, Simon Connell and Marius Tchonang

University of Johannesburg

E-mail: cookish@gmail.com

Abstract. The Mineral PET project aims to locate diamonds within Kimberlite by first irradiating the rock to induce beta decay in carbon atoms, then imaging the resultant positron emission signals. The presence of homogeneously distributed carbon forces us to go beyond simple counting techniques and rather create a 3D image of hotspots. The resolution of the image obtained is critically important for the location of small diamonds (of the order of 1mm). If the resolution is improved, less activation and/or fewer detectors are required to differentiate the diamond signal from the background.

While in theory the position and orientation of the detectors is known, each detector will differ slightly from the ideal values due to small misalignments, calibration differences etc. We therefore present attempts to improve position readouts inspired by techniques developed at CERN, specifically in the ATLAS detector of the LHC. Detector alignment at ATLAS uses a track-based algorithm that minimises a χ^2 value based on track-hit residuals. While the Mineral PET context is simplified by a vastly reduced number of detector elements, it is complicated by the fact that we are not interested in the path of the gamma rays that are actually detected, but rather the path of the diamond through the detectors, which can only be reconstructed from the combination of several detected events.

1. Introduction

The Mineral project images diamonds within coarsely crushed Kimberlite (± 10 cm rocks), without needing to bring diamonds to the surface. The technique uses Positron Emission Tomography (PET). The Kimberlite is irradiated, exciting the giant dipole resonance, to produce the unstable ^{11}C isotope. This beta decays, and the positron emission and annihilation leads to back-to-back 511 keV photons. These are detected in coincidence by planes of position sensitive detectors on either side of the passing rock. The image is then created by constructing a 3D voxel space with tubes drawn between pairs of detector events. The width is given by the position resolution of the detectors. This leads to a three dimensional density map of PET isotopes, with high activity indicated by the intersection of many tubes.

Only coincident events are recorded, and the utilisation of detectors and electronics with a very good time resolution ensures that the rate of random coincident events is low. There is still a true coincident background due to non-diamond positron annihilation events though. This is partly due to homogeneously distributed carbon throughout the Kimberlite; carbon makes 0.04% to 0.06% of Kimberlite by mass, depending on the type of Kimberlite. Secondly, oxygen will inevitably also be activated in the irradiation process, and it typically constitutes around

50% of Kimberlite. Luckily, ^{15}O has a half-life of 2 minutes compared to 20 minutes for ^{11}C .

The relative activity has been determined from experimental data taken at the Aarhus university synchrotron. The size of the 511 keV peak was found for data bins split according to successive time intervals. This was then fitted with a sum of exponential decays corresponding to each of the PET elements that are present, allowing us to find the relative concentrations of PET elements. Extrapolating back to the time of irradiation shows that the initial oxygen signal is 29 times higher than that of carbon. After 20 minutes therefore, diamond will have a PET activity between 2400 and 1600 times higher than that of the surrounding carbon and oxygen.

To be interesting in a real world context, the mineral PET technique would need to find a diamond as small as 1mm^3 . Given a 1mm diamond within a typical 10cm Kimberlite block, the non-diamond PET signal still outweighs that within the diamond by about 500 times. It is for this reason that we need to create a density image to look for hotspots, rather than rely on simple counting techniques. Given the current detector resolution of 5mm, and adding in the increase in resolution due to the interpolative effect of considering many different tubes, a well-aligned and calibrated system will easily differentiate a 1mm diamond from the background. Detector misalignments or inaccurate position calculation will however act to spread the diamond signal out into the noise. The higher the signal to noise ratio, the fewer events are required to identify the presence of a diamond. Good detector optimisation thus reduces the required irradiation beam current and/or the number of detectors, making it an important factor when analysing the feasibility of the system.

2. Alignment algorithms at ATLAS

The inspiration for our detector optimisation approach arises from the alignment procedure used in the ATLAS Transition Radiation Tracker (TRT). The approach used is an extension of the track fitting formalism [1]. One defines a global χ^2 as follows:

$$\chi^2 = \sum_i \left(\frac{m_i - h_i(x)}{\sigma_i} \right)^2 \quad (1)$$

where m_i is the set of all measured coordinates, x is a vector that parameterises all tracks under consideration and σ_i is the measurement error. h_i is the measurement model; $(m_i - h_i(x))$ is thus the difference for the i th coordinate between what was measured, and what the result should have been given parameters x . Finding the best track parameters means minimising χ^2 , typically by using a Newton-Raphson method to find where the first derivative is zero.

In the real world, the accuracy of measurements depends not only on the track parameters, but also on the alignment of the detectors that perform the measurement. x in (1) is thus replaced by a vector composed of both the track parameters x and calibration parameters α . This vector describes the position, rotation, etc. of each detector involved. One therefore simultaneously aligns the detectors and fits the particle tracks.

The problem described above can be expressed in matrix formalism as

$$\chi^2 = \sum_{\text{tracks}} r^T V^{-1} r \quad (2)$$

where r is the residual vector of differences between measured coordinates and our measurement model, $r(x, \alpha, m) = h(x) - m$. The $1/\sigma_i^2$ factor in (1) has been expressed more generally by the inverse of the covariance matrix, V^{-1} . In the context of ATLAS, there are a huge number of alignment and track parameters, and complex operations (particularly matrix inversion) are not computationally tractable with this number of dimensions. Work has therefore gone into

methods to minimise χ^2 without needing to perform the full matrix inversion; for instance splitting the problem into non-correlated components (see, for example, [2] and [3]).

There is an important difference between the typical detector alignment problem discussed above, and the Mineral PET alignment problem. The position of a positron source in Mineral PET is inferred from the intersection of randomly orientated tubes, rather than directly measured. A measurement model that tells us where hits will be detected cannot be defined, which means it is not possible to directly calculate a χ^2 based on residuals, as in (2). New measures of alignment accuracy must be therefore introduced, as discussed below.

3. Minimisation approach

3.1. Local parameters

“Local” parameters refer to variables that influence one detector at a time without requiring consideration of the correlations between detectors. The position readout from the resistive anode on the photomultiplier tubes is calculated in two dimensions by comparing the ratio of signals from opposite sides. This is not perfectly accurate. We thus introduce a series of corrections, which include: linearly expanding raw voltages, quadratically expanding x and y coordinates, adjusting the angle in polar coordinates, quadratically expanding the radius and linear scaling of x vs. y .

The scintillator crystals in the detectors are segmented into 5mm blocks, so we know exactly where the pixel should be. Minimisation of the local parameters can proceed for one detector at a time by comparing measured to actual pixel locations, without needing to worry about coincident hits. If the fitted position of the (i, j) th peak is \mathbf{x}_{ij} with error σ_{ij} , and the actual position for the (i, j) pixel is \mathbf{u}_{ij} , then a χ^2 score defined by

$$\chi_{\text{local}}^2 = \sum_{i,j} \frac{|\mathbf{x}_{ij} - \mathbf{u}_{ij}|^2}{\sigma_{ij}^2} \quad (3)$$

can be minimised for each detector. In practice, the pixels are similar, so to decrease computational load we set $\sigma_{ij} = 1$. Minimisation requires the following at every step: redraw the two dimensional position histogram with current parameters, find the peaks for each pixel and calculate the score from (3). Section 4.1 shows sample results.

3.2. Global parameters

The position and rotation of detectors cannot be optimised by considering only one detector at a time, as moving one detector changes the orientation of all tubes connecting it to other detectors. A strategy is therefore needed that is able to evaluate the performance of the system globally. To this end, data is taken with a point source. The alignment quality is determined by analysing the sharpness of the peak corresponding to the point source in voxel space. The initial approach was to fit the peak with a three dimensional Gaussian, of the form

$$f(x, y, z) = Ae^{-\left(\frac{(x-x_0)^2}{2\sigma_x^2} + \frac{(y-y_0)^2}{2\sigma_y^2} + \frac{(z-z_0)^2}{2\sigma_z^2}\right)} \quad (4)$$

Peak width is given by $\sigma_x\sigma_y\sigma_z$. This approach was found to not be practical because the peak shape often did not closely resemble a gaussian. Also, the sharpness of the peak can artificially improved, for example if detector positions are changed such that tubes are moved out of the fit window. This improves $\sigma_x\sigma_y\sigma_z$ despite decreasing alignment quality. Finally, requiring a full best fit computation for every score calculation is prohibitively slow computationally.

We therefore introduce a new method to quantify peak sharpness, S , defined as:

$$S = \sum_{(i,j,k) \in W} V(i, j, k)^2 \quad (5)$$

where $V(i, j, k)$ is a three-dimensional histogram that stores the sum of the tubes passing through the voxel labelled by coordinates i, j and k . The summation extends over a window W around the point source coordinates. By taking the square of the voxel counts, the score is increased both by sharpening the peak corresponding to the point source, as well as ensuring that as many tubes as possible pass through the fitting window W . In a sense the score can be thought of as a χ^2 fit for a flat background, and a good, sharp peak is one that fits this flat background as badly as possible. Minimisation proceeds by evaluating the inverse, S^{-1} .

Some non-sensical solutions will give rise to minima, such as all detectors placed exactly on top of each other. The more extreme examples of this type be avoided by sensible parameter constraints. In general however, it is important to consider more than one position of the point source at a time. The first step in this direction is to send the point source on a constant velocity track through the detectors. The positions of all detector hits are extrapolated backwards based on their timestamp, allowing the creation of a static voxel space as before. One individual track is still limited in its positional coverage however. The next step is thus to simultaneously consider several tracks passing through the voxel space. The collective score S_{tot} is defined as the sum of the individual track scores. Data is then taken by sending a point source through the detectors several times along different tracks. Because the tracks take place at different times, once the backwards extrapolation for velocity has been done, each track will create a point source peak at a different position in voxel space. A measurement window W can therefore be defined for each track, allowing simultaneous minimisation without interference.

One problem with the algorithm as outlined so far is that it is highly dependent on the accuracy of the tracks passing through the detectors. Taking a leaf from the ATLAS algorithm described in section 2, we therefore combine the track fitting and alignment algorithms. This is done by adding 6 additional parameters to the minimisation problem for each track in the data set. These parameters describe the initial position and velocity of the track. That starting time of the track is treated as a known; any inaccuracies in this can be accommodated by tweaking the initial position.

With two angles and three position coordinates for each detector, together with the track parameters, minimisation must therefore proceed in a $80 + 6(N_{\text{tracks}})$ dimensional space. This is further compounded by the fact that function is stochastic, and therefore not necessarily smooth, making the numeric determination of quantities such as derivatives unreliable. Finding a global minimum in such an environment is a daunting task for a numerical algorithm. We do not necessarily have to find the single, unique, global minimum to have success however. Any point that is found with a lower score than the starting point represents an improvement in alignment, even if it is not optimal.

We can improve the chances of success by aligning the parameters more closely with the physical situation. Consider the likely scenario of a mispositioned detector bank. To discover this, the minimisation algorithm would need to try a step where it simultaneously moved all detectors in the bank in one direction. To make this easier, we introduce additional parameters that capture the collective position and rotation of each detector bank. The minimisation then occurs hierarchically. First, track velocity and position parameters are minimised. This should be reasonably accurate given broadly correct starting values for detector positions. Next, the collective detector bank parameters are optimised. The individual detectors are optimised last.

4. Results

4.1. Single detector minimisation example

As proof of principle, the minimisation algorithms from the Minuit package [4] were applied to a single detector data set. Minuit is not the ideal package, as it is designed to find local minima. With some hand holding though, significant improvements were achieved. The Simplex method within Minuit was more effective in general, as it does not require well behaved derivatives

as opposed to the default Migrad algorithm. Figure 1 shows a two dimensional histogram of positions recorded by a single detector. The minimisation was able to improve χ_{local}^2 from 2.33 to 0.95. Visually, the “stretching” effect in the middle has been markedly reduced, with a more even pixel distribution. The Spectrum2 package was used for two dimensional peak finding [5].

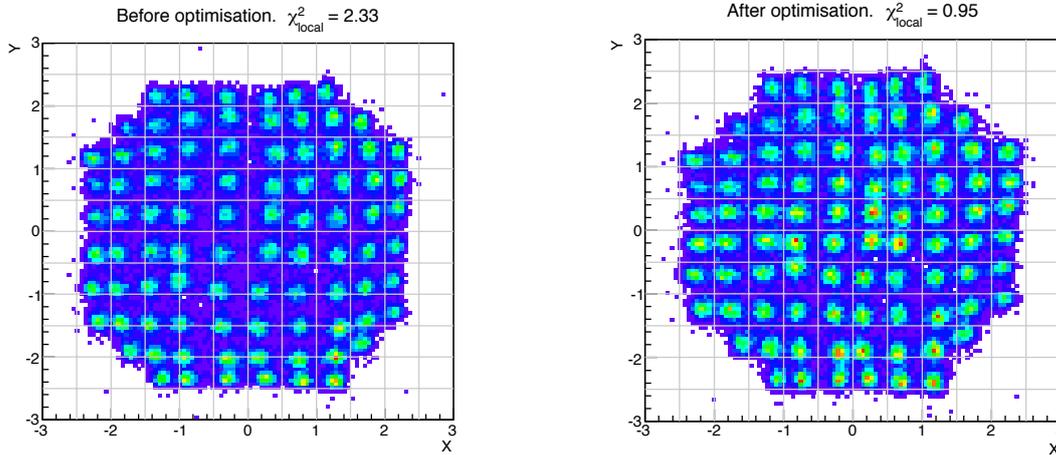


Figure 1. Single detector position graph before and after parameter optimisation. Warmer colours indicate higher counts. The grey grid lines show pixel boundaries.

4.2. Coincident minimisation

The full minimisation procedure has not yet been completed. Some preliminary results are shown below. A full image was taken with all 16 detectors. Our best measured parameters give the score from (5) to be $S = 17259$. The recorded position of one detector was then purposefully moved 2cm along the $+x$ axis. The resulting PET image is shown in figure 2, with $S = 13938$.

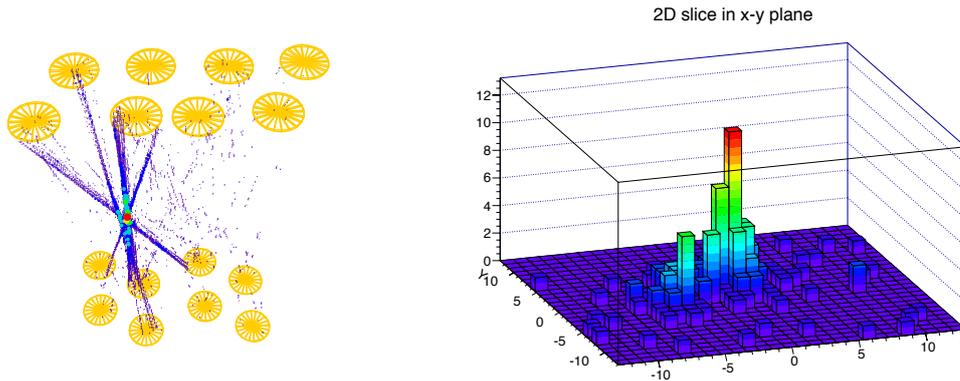


Figure 2. Left: 3D PET image showing voxel space after one detector was misaligned by 2cm. A threshold of 2 was used to reduce noise. Right: $x - y$ plane of the fitting window from (5).

The minimisation algorithm was then run to optimise the parameters of the moved detector. Figure 3 shows the resulting image, which has an improved score of 17743. One can see that the tubes are now properly intersecting at a point, creating a better resolved point source image. The minimisation procedure found the following changes: $x \rightarrow x - 2.05$, $y \rightarrow y - 0.48$, $z \rightarrow z + 0.57$. It also rotated the detector by 0.3 degrees. The fact that the score even exceeds the initial 17259 shows that the algorithm found values that were even better than the original “correct” ones.

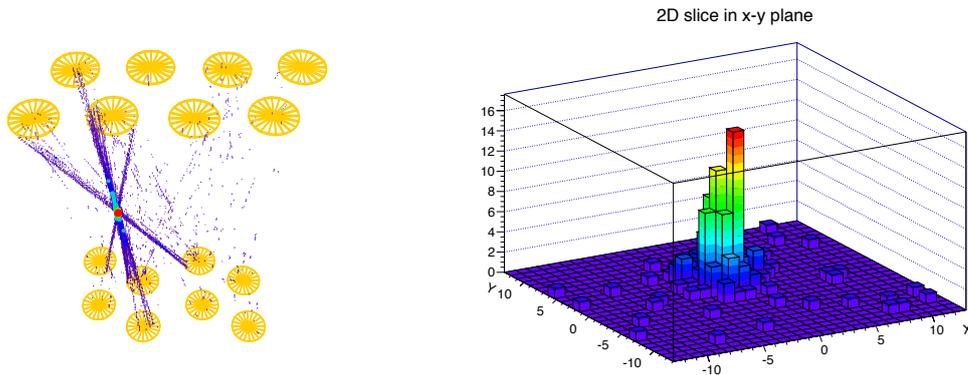


Figure 3. PET image after optimisation.

5. Conclusions and discussion

The hierarchical approach to parameter optimisation has produced some promising preliminary results. The major difficulty encountered so far is the behaviour of the minimisation algorithms. Gradient based algorithms fail badly, because the stochastic, uneven nature of the parameter space means that derivatives cannot reliably be calculated. Simplex based methods have produced more success, but need some hand-holding to ensure that the step size is small enough to jump over local minima and approach the true minimum. Future research will investigate specialised global (as opposed to local) minimisation algorithms, for example the method proposed in [6] that successively subdivides simplexes that cover the searchable region.

The second major obstacle is the large number of dimensions, which makes it intractable to minimise all parameters simultaneously. The hierarchical approach mitigates this to some extent by starting with global parameters that describe collective detector misalignments first.

The last obstacle is computational ability. Each function evaluation requires several seconds to compute, as a 3D histogram must be populated and analysed at every point. This is not a problem if only a handful of parameters are minimised at a time, but many function evaluations are needed for a minimisation algorithm to hope to properly search a higher dimensional space. This might require parallelisation of the code to run on a cluster.

In conclusion, it seems that reliably finding the unique global minimum that optimally aligns all detectors is a daunting task. Despite this, the method has already provided good improvements on our initial parameters, and promises to be a valuable technique to improve calibration even if the optimal global minimum is not found. We would like to extend our investigation to the optimisation of PET in other contexts, such as medical PET and positron emission particle tracking, using a moving point source based calibration cycle approach.

References

- [1] Bocci A and Hulsbergen W 2009 *ATL-INDET-PUB-2007-009*
- [2] Goetfert T 2006 *Iterative local 2 alignment algorithm for the ATLAS Pixel detector* Ph.D. thesis Julius-Maximilians-University Wurzburg and Max Planck Institute for Physics
- [3] Hartel R 2005 *Iterative local 2 alignment algorithm for the ATLAS SCT detector* Ph.D. thesis Technical University Munchen and Max Planck Institute for Physics
- [4] James F and Roos M 1975 *Comput.Phys.Commun.* **10** 343–367
- [5] Morhac M *et al.* 2000 *Nuclear Instruments and Methods in Research Physics A* **443** 108–25
- [6] Žilinskas A and Žilinskas J 2011 *Mathematical Modelling and Analysis* **3** **16** 451–460