

# Automated Track Recognition and Event Reconstruction in Nuclear Emulsion

P. Deines-Jones<sup>2</sup>, M.L. Cherry<sup>2</sup>, A. Dąbrowska<sup>1</sup>, R. Holynski<sup>1</sup>, W.V. Jones<sup>2</sup>, E.D. Kolganova<sup>4</sup>, D. Kudzia<sup>1</sup>, B.S. Nilsen<sup>2</sup>, A. Olszewski<sup>1</sup>, E.A. Pozharova<sup>4</sup>, K. Sengupta<sup>2†</sup>, M. Szarska<sup>1</sup>, A. Trzupek<sup>1</sup>, C.J. Waddington<sup>3</sup>, J.P. Wefel<sup>2</sup>, B. Wilczyńska<sup>1</sup>, H. Wilczynski<sup>1</sup>, W. Wolter<sup>1</sup>, B. Wosiek<sup>1</sup>, and K. Wozniak<sup>1</sup>

1. Institute of Nuclear Physics, Krakow, Poland
2. Louisiana State University, Baton Rouge, LA USA
3. University of Minnesota, Minneapolis, MN USA
4. Inst. of Theoretical and Experimental Physics, Moscow, Russia

## Abstract

The major advantages of nuclear emulsion for detecting charged particles are its submicron position resolution and sensitivity to minimum ionizing particles. These must be balanced, however, against the difficult manual microscope measurement by skilled observers required for the analysis. We have developed an automated system to acquire and analyze the microscope images from emulsion chambers. Each emulsion plate is analyzed independently, allowing coincidence techniques to be used in order to reject background and estimate error rates. The system has been used to analyze a sample of high-multiplicity Pb-Pb interactions (charged particle multiplicities  $\sim 1100$ ) produced by the 158 GeV/c per nucleon  $^{208}\text{Pb}$  beam at CERN. Automatically reconstructed track lists agree with our best manual measurements to 3%. We describe the image analysis and track reconstruction techniques, and discuss the measurement and reconstruction uncertainties.

## 1 Introduction

Nuclear emulsion is an excellent charged particle detector. It combines sensitivity to minimum ionizing particles (MIPs) with spatial resolution superior to the best electronic techniques available. This combination accounts for emulsion's usefulness in high energy cosmic ray experiments [1], neutrino oscillation searches [2], and analyses of high multiplicity heavy ion interactions [3]. Unfortunately, it has proven difficult to analyze emulsion in a systematic and automatic way, although attempts to do so date back at least to the 1950's [4]. Instead, measurement has

been a slow, manual task requiring a high degree of training, a fact which has limited both the number of analyzed events and the study of systematic errors in individual datasets. Automatic charge measurement in emulsion has long been possible in certain circumstances [5], and semi-automatic “bookkeeping” aids have been employed for some time [6, 7, 8]. But track counting and measurement in emulsion has remained a labor-intensive task.

Ironically, this difficulty is a consequence of emulsion’s advantages – high spatial resolution and sensitivity to MIPs – which make automatic track detection computationally challenging. In performing manual measurements, one continually adjusts the microscope’s focus slightly and looks for tracks that persist from the top of the emulsion to the bottom. To imitate this behavior, an automatic system must acquire many images, each at a slightly different focus depth, and the image analysis software must search for persistently dark paths through the resulting three-dimensional “focus sequence” of image frames. For this reason, large quantities of imaging data must be acquired and processed. The analysis routines must efficiently detect tracks yet reject the background from knock-on electrons, secondary particle production, etc. Until recently [9, 10], this data acquisition and analysis was impractical.

The Krakow-Louisiana-Minnesota-Moscow collaboration (KLMM) has developed a system which automatically measures and reconstructs nuclear interactions in emulsion “chambers”, in which thin emulsion plates are exposed perpendicular to the beam. The system identifies the particles that emanate from a common vertex, efficiently rejects background tracks, measures the track space angles, and provides a rough charge assignment which distinguishes minimum ionizing tracks from heavier fragments. The overall reconstruction accuracy is 97% or better. We have used this system to analyze a set of 40 semi-central 158 GeV/c per nucleon Pb-Pb events with multiplicities ranging from about 600 to 1700, demonstrating

for the first time the scientific utility of such a system [3, 11].

Section 2 describes the KLMM Pb-Pb emulsion chamber experiment. The image acquisition is discussed in Section 3, and Sections 4 and 5 cover the image analysis and track reconstruction procedures.

## 2 Chamber Design and Exposure

KLMM exposed 32 emulsion chambers to the 158 GeV/c  $^{208}\text{Pb}$  beam accelerated at CERN in December 1994 (CERN experiment EMU-13). The emulsion chambers consist of Pb foil targets and emulsion plates oriented perpendicular to the beam (Fig. 1). In contrast to stacks of emulsion pellicles oriented parallel to the beam, emulsion chambers can use pure targets, allowing the study of symmetric Pb-Pb collisions free from possible target selection biases. Angular measurements in chambers (as opposed to stacks) are relatively free from systematic uncertainties due to emulsion shrinkage and distortion. Unlike emulsion stacks, however, chambers measure only forward-going tracks.

By using emulsion only to sample the path of the track, chambers present very little grammage either to the incident beam or to produced particles, thus reducing electron pair production and secondary interactions. However, the relatively short path length in emulsion allows only a rough charge assessment in the chamber itself. Three slanted pellicles, the most downstream elements of the chamber, can be used to assign more precise charges to the spectator fragments.

Each emulsion plate consists of a 200  $\mu\text{m}$  thick acrylic base coated with a 55  $\mu\text{m}$  Fuji ET7B emulsion layer on each side. (Each plate consists of only  $\sim 0.06 \text{ g/cm}^2$  of material. Most tracks are fully measured before they pass through 4 such plates.) The Pb targets and plates have faces measuring 10 cm  $\times$  5 cm. Each chamber holds either 3 or 4 100  $\mu\text{m}$  thick lead target foils. Milled Rohacell spacers 1.00 cm

and 1.50 cm thick provide accurately known longitudinal distances necessary for reconstruction of space angles from plate position measurements.

The exposure of the chambers to the 158 GeV/c  $^{208}\text{Pb}$  beam resulted in an average of  $\sim 350$  primary  $^{208}\text{Pb}$  ions/cm<sup>2</sup> across the face of the chambers, concentrated in three  $1.5 \times 2$  cm<sup>2</sup> beam spots. This density is small enough to ensure a low delta-ray background and to keep the data cuts due to interactions occurring too close to a non-interacting primary to an acceptably low level.

### 3 Event Acquisition

Event reconstruction through the analysis of microscope images is done in three stages: event acquisition, identification of track candidates in the digitized images (“image analysis”), and reconstruction of the event from the lists of track candidates. The entire processing chain is shown in Fig. 2. The data-taking phase consists of:

- scanning, which locates and selects candidate events for study,
- tracing, which ensures the interaction occurred in a Pb target,
- image acquisition, which records microscope fields around the event in several plates, each spanning a different range in opening angle.

#### 3.1 Scanning and Tracing

The emulsion plates directly below each target are visually scanned at low magnification (200x) for events. (The scanning procedure for the initial sample of 40 semicentral Pb-Pb events is described in [3].) After the initial scanning selections are made, each event is examined in all the plates upstream of the interaction and rejected if the primary was noticeably less ionizing (approximately 5 charge units) than nearby Pb tracks or if the primary suffered an additional observable

interaction. The plates immediately upstream and downstream of the target are examined, and events which occur in emulsion rather than the Pb target are rejected. Events with nearby (less than  $60\text{ }\mu\text{m}$ ) non-interacting primaries which might obscure secondary tracks are also rejected. The plate position of each scanned event is recorded to  $\pm 0.5\text{ mm}$ . A low-magnification locator image is also recorded, with the event in the center of the image. The locator image includes nearby non-interacting primaries, and makes it possible to determine the event position to  $\pm 5\text{ }\mu\text{m}$  relative to the nearby non-interacting tracks.

### 3.2 Image Acquisition

To digitize the emulsion images for event reconstruction, we have constructed several microscopy systems equipped with PC-controlled stages and CCD cameras (Fig. 3). In the usual “high-power” mode of operation, a 100x microscope objective together with a 0.45x coupling lens yields a useful image which is  $108\text{ }\mu\text{m} \times 140\text{ }\mu\text{m}$ , and which has about a  $\sim 2\text{ }\mu\text{m}$  depth of field. In the typical “low-power” mode, a 6x objective gives a  $2.3\text{ mm} \times 1.8\text{ mm}$  field of view, with a depth of about  $200\text{ }\mu\text{m}$ . The digitized pictures are 512 pixels  $\times$  480 pixels  $\times$  8 bits. The microscope stage is equipped with stepping motors and linear optical encoders on all three axes. It can be stepped under software control in  $1\text{ }\mu\text{m}$  steps in three directions, or it can be operated manually.

To acquire a focus sequence, the emulsion plate to be imaged is manually registered to  $\pm 5\text{ }\mu\text{m}$  with respect to the event’s locator image with the aid of a blink comparator, which can switch between the “live” camera view and the locator image. During acquisition of a focus sequence, the stage is controlled by the image acquisition program. This program monitors the CCD image and begins acquisition starting at the upper surface of the emulsion. It then steps the focus vertically in  $0.8\text{ }\mu\text{m}$  steps until it finds the lower surface, at which time

it terminates acquisition and writes the focus sequence to a file. Surfaces are detected by subtracting consecutive frames and finding the largest absolute residual in a selected window. If  $I(x, y, z)$  is the image brightness at a pixel located at coordinates  $(x, y, z)$  and  $\Delta z = 0.8\mu\text{m}$  is the focus step size between consecutive frames, the focus signal  $F$  is

$$F = \max_{(x,y)} |I(x, y, z + \Delta z) - I(x, y, z)| \quad (1)$$

resulting in a focus signal like the example in Fig. 4. (This subtraction works because there are almost always at least a few grains in focus when the microscope is focused in the emulsion. Moving the focus  $0.8\mu\text{m}$  makes these in-focus grains significantly more blurry and brighter.) To avoid triggering on objects outside the emulsion, such as dust or air bubbles in the immersion oil (used to optically couple the microscope objective to the emulsion) this calculation is performed in four separate windows; the extreme values are discarded and the second highest value is kept. The focus signal is digitally filtered to debounce the transition and the result compared to a preset threshold to determine whether the microscope is focused inside or outside the emulsion [11]. Depending upon the exact emulsion thickness, approximately 20 frames are acquired in each focus sequence.<sup>1</sup> The determination of the emulsion thickness is repeatable to  $\pm 1\mu\text{m}$ .

## 4 Image Analysis

Image analysis begins with a focus sequence of images and ends with a list of track candidates and their coordinates for that sequence's field of view. The analysis must efficiently discriminate secondary tracks (the signal) from the various backgrounds. It must do it quickly, and therefore simply; since 15–20 such fields of view

---

<sup>1</sup>Note that  $20 \times 0.8\mu\text{m} = 16\mu\text{m}$  is a typical emulsion thickness *after* development, and is substantially less than the nominal  $55\mu\text{m}$  pre-development thickness quoted above.

are analyzed to reconstruct one typical event, speed is an issue if the system is to be practical. To develop the analysis, the ideas of emulsion “signal” and “background” need to be articulated precisely enough so that they can be translated into computer code. The software might be written to hunt for individual grains, and then assemble them into tracks; it might treat the tracks themselves as primitive objects; or it might recognize an interaction vertex as a “gestalt”. We have settled on the last strategy, which provides excellent signal-background separation while at the same time being computationally practical.

Secondary (i.e., highly relativistic) tracks in emulsion have a straight, ray-like appearance. Depending on their charge and angle of inclination, they appear either as a series of distinct grains, randomly distributed along the track, or a more or less solid track of ionization, perhaps accompanied by occasional delta rays. (A track which is viewed almost end-on is not resolved into distinct grains.) In any case, a minimum ionizing particle produces on average one developed grain every  $3.5 \mu\text{m}$  along its path, yielding  $16 \pm 4$  grains in  $55 \mu\text{m}$  of emulsion. The individual grains appear at high power as small regions ( $\sim 0.5 \mu\text{m}$ ) which are 40–70% as bright as their surrounding neighborhood. Small angle Coulomb scattering is negligible in 55 microns of emulsion for even the lowest energy produced particles. Secondary interactions are quite rare; the pion nuclear m.f.p. in emulsion is 35 cm. The geometry of secondary tracks is therefore simple: to a very good approximation, they are straight tracks that point back to a common vertex.

The physical backgrounds can be grouped into two categories. In the first group are “random tracks,” which are straight but are not associated with the event under study. The only way to distinguish these real but unrelated tracks from those which are created by the interaction is by confirming whether or not they point back to the vertex. The other kind of background tracks are delta rays, which scatter significantly in a single emulsion layer, and deposit more ionization energy

in emulsion than more massive MIPs. Heavy ion beam tracks copiously produce long-range delta rays, and some of these escape the emulsion plate in which they were produced, giving rise to a fairly uniform distribution of delta rays on top of the local distribution surrounding each beam track.

Among the instrumental backgrounds are “chemical fog,” consisting of developed grains which are not associated with any ionizing track, but are an artifact of the development process. Emulsion surface defects may also be prominent enough to cause problems, especially if the emulsion is thin.

The last kind of background, shadowing, is not strictly a background at all; rather, it is an instrumental effect. In ordinary transmitted light microscopes, the light passes through the entire two-sided emulsion plate before reaching the eye or CCD. Thus, the objects near the plane of focus are not uniformly illuminated, but are shadowed by out-of-focus objects below (and above) them. The magnitude of the darkening of the field due to shadowing is of the same order of magnitude as the darkness of the grains themselves.

The natures of the signal and backgrounds give us some clues about how a successful track recognition algorithm should work. Because the individual grains in a track are not always resolved, and also because many or most grains are not part of secondary tracks, it is reasonable to try to detect the entire track rather than the grains of which it is composed. We could therefore operationally define a track to be a straight path through the emulsion which has some minimum average darkness. This criterion excludes chemical fog, since single grains contribute only a small amount of darkness to any path through the emulsion. For similar reasons, it also discriminates against delta rays. However, because of their scattering, delta rays develop more grains than other MIPs, and sometimes mimic real secondaries, especially if they are energetic enough to follow more or less straight paths for 20–30  $\mu\text{m}$ . We therefore need a second criterion – that the dark path be small



and compact in the transverse direction in order to ensure that the particle that produced the path did not scatter. We accomplish this by demanding that the path be darker than similar paths in its local neighborhood [11]. This criterion also solves the shadowing problem, since we measure the darkness of tracks not in terms of the intensity of light incident on the plate, but relative to the brightness in their immediate neighborhood.

Finally, we require that all selected tracks point back to a common vertex. It is important to realize that the vertex point one sees in the emulsion does *not* correspond directly to the event coordinates of the interaction point due to emulsion shrinkage and distortion, as well as to the uncertainty in the measurement reference system. We need to identify tracks which point back to this “apparent vertex” whose position is known *a priori* to  $\pm 5 \mu\text{m}$  in the transverse direction and to within 5–50% in the longitudinal direction, relative to the center of the microscope field of view. (The larger value applies to the plate closest to the target.) Because of the uncertainty in the apparent vertex position we need to modify the vertex criterion slightly: we demand that all secondary tracks point to a common apparent vertex *whose position will have to be determined as we search for tracks*.

This new vertex requirement brings us to the conclusion promised above: the software will search for an apparent vertex, rather than individual tracks. Once the vertex is found, the individual tracks of which it is composed can then be identified and characterized.

Fig. 5 illustrates how the vertex finding is done. For each trial vertex, the intensity in individual frames is averaged along paths radiating from the trial vertex. This produces a processed image which can be thought of as what the emulsion would look like from the standpoint of the trial vertex. Tracks passing through the trial vertex appear as dark spots, while isolated grains, coincidental

tracks, and delta rays appear washed out. The vertex finder evaluates trial vertices and searches for the one with the maximum number of small dark spots. Fig. 6 compares such a processed or “accumulated” image to one of its constituent frames. To count the number of tracks in the accumulated image, the vertex finder first high-pass filters the image, which imposes the compactness criterion by removing large (diameters greater than  $1\ \mu\text{m}$  or so) objects, and also removes the shadowing bias. The pixel darknesses in the resulting image are then compared to a threshold (Fig. 7), producing yet another image in which the dark pixels are turned on and the bright pixels turned off. Each distinct cluster of dark pixels is counted as a candidate track, and the optimization routine in the vertex finder maximizes the number of clusters [12] to determine the best apparent vertex and produce the final accumulated image, which is stored for further analysis [11].

The filtering routine in the vertex finder is optimized for speed, since it is not necessary to find every single track to accurately determine the apparent vertex position. The final accumulated image is therefore handed off to a second-stage image analysis routine which performs essentially the same analysis but in a more careful manner. Each resulting cluster is centroided to measure the track position. In addition, each track’s darkness is measured by comparing the mean brightness of pixels around the track centroid to the pixel brightness off-track. The resulting list of track positions and darknesses of each candidate track is saved for later submission to the plate fitting and reconstruction routines.

To give a qualitative idea of the accuracy of the image analysis, track candidates from the same event measured in two different emulsions are compared in Fig. 8. The measurements in the upstream emulsion have been scaled and shifted to overlap with the downstream layer. The correspondence is quite good, but, as expected, there are some candidates in one emulsion that do not appear in the other. Either a real track was missed in one layer, or background was incorrectly

identified as a secondary track. One can see from this comparison that it is possible to clean the candidate track lists by comparing consecutive emulsions. The vertex finder analyzes each field of view independently, and this allows us to use coincidence techniques both to clean the track list and to systematically estimate backgrounds and efficiencies. To do this, we must first assemble all of the individual emulsion track lists into a single list for the entire event. This is the subject of the next section.

## 5 Reconstruction

The image analysis produces track lists from each individual measured emulsion. The reconstruction routine must then search all the emulsions for the individual measurements along each track and join them together to form a single track list. Reconstruction entails precisely determining the emulsion positions relative to one another and to the vertex, and then comparing the individual measurements in all the plates to find the real tracks and reject the background.

When comparing track measurements in two different plates, it frequently is not straightforward to match pairs of measurements of the same track. In order to connect the measurement pairs, one must know the relative positions of the emulsions and the vertex. However, the uncertainties in these positions, which have been determined from local non-interacting primaries ( $\pm 5 \mu\text{m}$ ) and knowledge of the chamber structure ( $\pm \sim 300 \mu\text{m}$ ), are far too large for positive assignment of individual measurements to particular tracks. Fig. 9 schematically illustrates the plate alignment problem. Almost all produced particles emitted from the interaction have virtually straight trajectories [Fig. 9(a)]. After disassembling the chamber for development and image acquisition, we have imprecise knowledge of the plate and vertex positions, which makes track reconstruction ambiguous

[Fig. 9(b)].

In principle, the apparent vertices provide information about the plate positions, but using this information for plate alignment works poorly in practice, mainly for two reasons. A small amount of linear emulsion distortion can shift the apparent vertex horizontally many microns. In addition, precise knowledge of the emulsion shrinkage factor, a function of relative humidity and temperature, is required and entails careful manual measurement of every plate at the time of image acquisition. Instead, the plate alignment is done using pattern matching software [Fig. 9(c)]. As the figure shows, the pattern matching determines positions relative to the vertex up to a transverse shift (i.e., an uncertainty in the direction of the event axis) and a longitudinal scale. The transverse ambiguities are removed by assuming the event axis is parallel to the local non-interacting primaries (LPs), and the longitudinal scale is determined using the fiducial spacer. Once this information is incorporated, the original event geometry is reconstructed [Fig. 9(d)].

The pattern matching algorithm aligns a pair of emulsions by shifting the upstream emulsion measurements with respect to the downstream points by an offset  $(\Delta x, \Delta y)$  and by scaling the upstream measurements by a factor  $m$  in order to maximize the overlap between the two emulsions. To characterize the quality of the overlap, the figure of merit  $S$  that is maximized is

$$S = \sum_{i=1}^{N_{DS}} e^{-(d_{(nn)i}/\rho_0)^2}, \quad (2)$$

where  $N_{DS}$  is the number of tracks in the downstream side,  $d_{(nn)i}$  is the distance between downstream track  $i$  and its nearest neighbor in the upstream emulsion, and  $\rho_0$  is set to  $1.0 \mu\text{m}$ . For close pairs ( $d_{(nn)} \ll 1.0 \mu\text{m}$ ), the individual exponential terms approach

$$1 - (d_{(nn)i}/\rho_0)^2, \quad (3)$$

and  $S$  is a measure of the sum of the squares of distances between nearest neighbors. The exponent discounts tracks whose nearest neighbors are more than  $1.0 \mu\text{m}$  away, as these are likely to be spurious measurements. This fitting procedure is performed in a pairwise fashion, starting with the most downstream pair of emulsions and chaining up to the most upstream. For example, the downstream side of plate 5 is fitted to the upstream side of plate 5, which is fitted to the downstream side of plate 4, which is fitted to the upstream side of plate 4, etc. Every matched emulsion pair is plotted (e.g., Fig. 8) for visual inspection to confirm the fits. The fitted longitudinal positions of the emulsion plates are also compared with the known chamber structure to check for gross fitting errors.

Once the plate matching is complete, direction vectors which cluster together are assigned to each measurement, and the individual measurements can be grouped together into track candidates. The direction vector to a point from the vertex is characterized by  $(x_{ref}, y_{ref})$ , the point at which the trajectory intersects an arbitrary reference plane parallel to the plates at distance  $z_{ref}$  from the vertex. The direction vector is related to the space angles through

$$\begin{aligned}\tan \theta &= \sqrt{x_{ref}^2 + y_{ref}^2}/z_{ref}, \\ \tan \phi &= y_{ref}/x_{ref}.\end{aligned}\tag{4}$$

Qualitatively, a track candidate is a cluster of individual measurements with similar values of  $x_{ref}$  and  $y_{ref}$  (within  $\sim 1.0 \mu\text{m}$ ). A list of clusters is generated according to criteria which are unrestrictive enough to include almost all real tracks, and also some spurious candidates. The main requirements for a candidate to be considered a confirmed track are:

- A candidate must be measured in at least two emulsions. This coincidence requirement efficiently discriminates against residual background.

- The candidate cannot be missed in more than two consecutive emulsions. This requirement cuts accidental coincidences, and also tracks which do not point precisely back to the vertex.

In addition, there are further tests against low-energy tracks and tertiary electron-positron pairs, and against spurious close pairs (caused by one false measurement close to a real track). These selections are described in the Appendix.

## 6 Results

To test the image analysis and reconstruction techniques, two automatically measured Pb-Pb events, with multiplicities  $\sim 670$  and  $1300$ , were checked using the semi-automatic microscope system at INP in Krakow. The INP system uses a CCD camera to display the microscope image on a monitor, allowing tracks to be measured by the operator using a mouse and cursor. These track measurements are stored in a computer file as they are taken. The events are reconstructed manually as the measurements are taken. A similar test was performed using two  $10.6$  GeV per nucleon Au-Au events with multiplicities of about  $120$ . The Au chambers provide a more rigorous test of background rejection (these chambers were exposed to  $\sim 3000 \text{ cm}^{-2}$  primaries), but the Pb events have larger multiplicities, the secondary tracks more densely populate the plates, and there are many more plates applied in the reconstruction ( $\sim 12$  for Pb, compared to 3 or 4 for Au). In both cases, the two methods agree on 97% of the reconstructed tracks. The distribution of discrepancies is uniform in azimuthal angle within statistical uncertainties, and the opening angle distribution of the discrepancies is consistent with a flat detection efficiency in the region of full acceptance ( $\eta > 2.6$  or  $\theta < 0.15 \text{ rad}$ ).

The main measurement biases can be estimated independently of this comparison. The coincidence technique allows the efficiency with which tracks are

detected to be estimated for every event, plate-by-plate. After the event is reconstructed, each track is examined for “missing” measurements, i.e., emulsions in which the track could have been detected (because it was well-separated from other tracks) but was not seen. A record of misses is kept for every event and every plate. Similarly, the “singles” background is estimated from the number of measurements which are not used in reconstructed tracks, and thus are presumed to be background. Fig. 10 shows these diagnostics for a sample event. Both the miss rate and the singles rate are roughly independent of the track density. The instrumental resolution of close pairs (Fig. 11) is  $1.0 \mu\text{m}$ , causing very close pairs of tracks to be undercounted. We estimate this bias to be 3% of the track count for these high multiplicity Pb-Pb events. Combining the biases calculated from the measured efficiencies, singles rates, and pair resolution (c.f. Appendix), the overall track counting errors are estimated to be 3% of the track count, in good agreement with the comparison to the semi-automatic measurements.

Having established that the reconstruction procedure produces “clean” track lists at the level of  $\pm 3\%$ , we can examine the accuracy of the angular measurements. Fig. 12 shows the standard deviation of individual measurements about their fitted tracks. The mean standard deviation is  $0.14 \mu\text{m}$  in both the x and y-directions, and the transverse measurement uncertainty is therefore  $0.20 \mu\text{m}$ . Since the field of view is about  $50 \mu\text{m}$  from center to perimeter, the track opening angles  $\theta$  are determined to  $\sim 0.4\%$ , corresponding to an uncertainty  $\delta\eta \approx 0.005$  in the pseudorapidity  $\eta = -\ln \tan \frac{\theta}{2}$ . A systematic uncertainty in the transverse positions derives from the absolute determination of the event axis. Typically, the last measured plate is 3.3 cm downstream. This results in a typical systematic uncertainty in the azimuthal angle  $\theta$  of  $5 \mu\text{m}/3.3 \text{ cm} = 0.15 \text{ mrad}$  in the absolute positioning of the event with respect to the reference system. The uncertainty in the longitudinal track positions has a statistical component which is greatest at

large angles but does not exceed 1%, and an estimated 1% systematic component due to uncertainty in the fiducial spacer thickness. The overall uncertainty in the pseudorapidity ranges from  $\sim 0.01$  at small  $\eta$  to 0.03 at  $\eta = 6$ . The value of the pseudorapidity loses significance beyond  $\eta = 9$ . Fig. 13 shows a typical pseudorapidity distribution, and the average of 40 high multiplicity pseudorapidity distributions.

Besides a track's space angle, its other main property is its darkness. Fig. 14 shows the darkness distribution near the interaction axis as a function of opening angle for 40 semi-central events. Most of the tracks within 5 mrad of the axis are minimum ionizing, but a more heavily ionizing component can also be observed, corresponding to spectator alphas and heavier fragments.

Using the automatic system, a single operator can measure and reconstruct events several times faster than previously possible. Because the analysis can be performed in parallel on several machines simultaneously, the measurement "bottleneck" is the image acquisition. With the current setup, a chamber with 20 events can be digitized in 3-5 days, and the analysis can be started while data from the next chamber is being acquired. A single Pb event is processed on a 166 MHz Pentium in  $\sim 4$  hours.

In summary, when the entire analysis chain from image analysis to track reconstruction is tested as a whole, the results agree well with careful manual measurements. Further, automatic measurement opens up new possibilities to rigorously understand counting systematics by providing consistent, detailed background and efficiency measures. This accomplishment augments one of emulsion's main strengths: the ability to characterize individual tracks. At the same time, automation ameliorates emulsion's chief weakness by making the measurement much faster, simpler, and significantly more repeatable and systematic.



## 7 Acknowledgements

This work was partially supported in the U.S. by the National Science Foundation (Grants PHY-921361 and INT-8913051 at LSU) and Department of Energy (DOE-FG02-89ER40528 at Minnesota), and in Poland by State Committee for Scientific Research grant 2P03B18409 and by Maria Skłodowska-Curie Fund II No. PAA/NSF-96-256. PD thanks the Louisiana State Board of Regents (LEQSF) under agreement NASA/LSU-91-96-01 and NASA/ LaSPACE under grant NGT-40039 for their support. Construction of the automated microscope system was funded by NASA (NAGW-3289 and NAGW-3560) at LSU. We very much appreciate the help of the CERN staff, A. Aranas, J. Dugas, and L. Wolf at LSU, and especially Prof. Y. Takahashi and his EMU-16 colleagues for their generous assistance.

<sup>†</sup> Current address: Horizon Comp., 5 Lincoln Hwy, Edison, NJ 08820.

## Figure Captions

Fig. 1: KLMM Pb chamber used at CERN. A chamber with three target modules is shown. Some of the chambers had four targets. The right-hand columns show details of the upstream chamber structure at 10 and 100 times the scale of the left column. The horizontal scale is arbitrary.

Fig. 2: Overview of the reconstruction analysis chain.

Fig. 3: The LSU automated microscopy system. The monitor can display the contents of either a stored frame buffer or the “live” digitized image, allowing it to be used as blink comparator between live and stored images, or between two stored images.

Fig. 4: Focus signal as a function of focus depth in emulsion. A large signal indicates the microscope is focused in the emulsion.

Fig. 5: Schematic illustration of the vertex finding process.

Fig. 6: (a) Event 20-06, approximately  $80\text{ }\mu\text{m}$  downstream of the interaction in the Pb foil. The field of view is  $140\text{ }\mu\text{m} \times 80\text{ }\mu\text{m}$ , and the depth of the field is about  $1\text{ }\mu\text{m}$ . The focal plane is  $4\text{ }\mu\text{m}$  into the emulsion. (b) Accumulated image constructed from (a) and 19 other frames in a focus sequence.

Fig. 7: Histogram of filtered image values on every darkness peak. The threshold is individually determined for every field of view based on the position of the background peak.

Fig. 8: Coincidence of measurements in two emulsions. In the upstream emulsion (crosses), which is about  $80\text{ }\mu\text{m}$  from the vertex, the tracks in the center of the field of view are not resolved. The downstream emulsion (squares) is 2.97 times farther downstream, and the tracks are resolved almost to the event axis. On the other hand, the upstream emulsion shows wide-angle tracks which have passed out of the field of view of the downstream emulsion. The units are arbitrary.

Fig. 9: The plate alignment problem. (a) Plate alignment prior to disassembling the chamber. The large circles represent a local noninteracting primary; smaller dots represent shower tracks. (b) Upon disassembly, plate registration is lost. (c) Pattern matching reconstructs plate positions up to an overall transverse shift and longitudinal scale. (d) Local noninteracting primaries determine the shift and the fiducial spacer determines the scale.

Fig. 10: Track count, singles rate, and measurement inefficiency for individual plates for Event 20-06. (In plate 16, 12 tracks are visible. Two of these are missed in emulsion 16U, giving rise to a large miss rate in this emulsion.)

Fig. 11: Pair resolution efficiency as a function of pair separation on the emulsion plate for a typical event. To calculate the resolution efficiency, two histograms of track separations are constructed for both fully resolved and unresolved measurements along a sample of reconstructed tracks. At each separation, the resolution efficiency is then  $(\text{resolved})/(\text{resolved} + \text{unresolved})$ .

Fig. 12: Standard deviations of individual measurements around their fitted track trajectories.

Fig. 13: Pseudorapidity distributions for a sample event (solid line), and the average pseudorapidity distribution of 40 high multiplicity events (dotted line) [3].

Fig. 14: Darkness distribution of tracks near the event axis. (a) Track darkness vs. opening angle. (b) Darkness distribution of all tracks in the forward 2 mrad cone.

Fig. 15: Features of track candidate clusters. Solid lines connect measurements in accepted tracks. Dotted lines represent rejected clusters. (a) A cluster consists of measurements with similar space angles. (b) Isolated single measurements are rejected. (c,d) Measurements can belong to more than one track. These examples are typical of close track pairs. (e,f) A close pair must be confirmed in the next layer. Track (e) is confirmed; (f) is not. (g) Gaps of one emulsion are allowed. (h)

Gaps of more than one emulsion are not allowed.

## References

- [1] T.H. Burnett *et al.*, Nucl. Instr. Methods A **498**, 583 (1986).
- [2] K. Winter, Proc. 16th Intl. Conf. on Neutrino Astrophysics (Nucl. Phys. B [Proc. Suppl.] **38**, 211) (1995).
- [3] P. Deines-Jones *et al.*, Phys. Rev. C **53**, 3044 (1996).
- [4] C.F. Powell, P.H. Fowler, and D.H. Perkins, *The Study of Elementary Particles by the Photographic Method*, London: Pergamon Press, 1959.
- [5] P.H. Fowler, Nucl. Instr. Methods **147**, 183 (1977).
- [6] A. Iyono *et al.*, Nucl. Instr. Methods B **52**, 98 (1990).
- [7] S. Garpman *et al.*, Nucl. Instr. Methods A **269**, 134 (1988).
- [8] E. Olson *et al.*, Proc. 23rd Intl. Cosmic Ray Conf. (Calgary) **4**, 718 (1993).
- [9] P. Deines-Jones *et al.*, Proc. 23rd Intl. Cosmic Ray Conf. (Calgary) **2**, 536 (1993).
- [10] A. Tawfik, private communication (1996); E. Ganssauge, invited talk at the 16<sup>th</sup> Intl. Conf. on Nuclear Tracks in Solids, Beijing, China (1992)
- [11] P. Deines-Jones, Ph.D. thesis, Louisiana State University (1996).
- [12] J. Hoshen and R. Kopelman, Phys. Rev. B **14**, 3438 (1976); H. Gould and J. Tobochnik, *An Introduction to Computer Simulation Methods* (Part II), Reading, MA: Addison-Wesley, 1988.
- [13] W.H. Press *et al.*, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge, MA: Cambridge Press, 1992.

## A Reconstruction Algorithms

### A.1 Plate Fitting

Essentially, the plate fitting program is a minimization routine which minimizes the sum of some measure of distance between nearest neighbors in two adjacent plates. The standard quantity to minimize is the sum of the squares of the distances, but in this case, the sum of the squares emphasizes distant pairs which are physically unrelated. Optimization routines based on summing the squares sometimes converge but are extremely unstable. Instead, we have chosen a function which acts like the distance squared for small distances, but contributes little to the sum at large distances (Eq. 2). The detailed behavior of this function appears to be irrelevant – we have obtained equally good results, for example, by substituting a lorentzian for the exponential. What is important is the behavior at small and large distances  $d_{nn}$ .

The figure of merit  $S$  is a function of the relative transverse plate shifts  $(\Delta x, \Delta y)$  and the ratio  $m$  of the distances of the downstream plate's distance to the vertex to the upstream plate's vertex distance. Because of the large number of measurements in both plates, the function frequently has several local minima. To correctly align the plates, the routine must find the global minimum. Like the vertex finder, the plate alignment routine has a first stage grid search followed by a conventional optimization routine (Powell's Method [13]). To do the grid search, each plate's track positions are binned into a two-dimensional array representing the field of view. The downstream plate is binned on a 2 micron square grid. The upstream coordinates are first transformed according to

$$\begin{aligned}x_T &= m(x - \Delta x), \\ y_T &= m(y - \Delta y)\end{aligned}\tag{5}$$

and then are also binned on a 2 micron grid. The quality of the overlap is found by calculating a match-to-miss ratio, where the number of “matches” is the number of array bins that contain tracks upstream and downstream, and the “misses” are the number of bins with an upstream track but no downstream track. (The mean nearest-neighbor track spacing is roughly  $5\text{ }\mu\text{m}$  near the edges of the plates. Therefore, the  $2\text{ }\mu\text{m}$  array elements are mostly 0 near the plate edges. The size of the array element must be small enough that there are a significant number of empty elements, but large enough that the calculation is performed quickly.) The grid search is performed on a  $2\text{ }\mu\text{m}$  transverse grid. The scale step size is chosen so that a single step changes the upstream positions  $(x_T, y_T)$  by no more than  $2\text{ }\mu\text{m}$ .

## A.2 Reconstruction

Reconstruction starts with the most downstream emulsion, and matches for each measurement are sought in the next most upstream emulsion. An upstream measurement is considered a possible mate if it falls within  $1\text{ }\mu\text{m}$  of the ray joining the vertex to the downstream measurement. If more than one measurement exists in the search radius, the nearest is selected. If no match is found in that emulsion, the next one is searched, and so on. This procedure allows an upstream measurement to be shared among two or more tracks, but ensures a branching structure, in which two tracks never rejoin downstream.

To be considered a confirmed track, each cluster must pass four tests, illustrated by the examples in Fig. 15:

**Coincidence** A track must be measured at least twice. This discriminates against residual background from the image processing stage. All the clusters in Fig. 15 except (b) meet this criterion.

**Dispersion** The RMS scatter around the fitted track trajectory must be less than  $1.0\ \mu\text{m}$ , corresponding to 5 standard deviations. This cuts tracks that do not point back accurately to the vertex, as well as low-energy tracks emitted by the struck target.

**Accidental background** This tests for random tracks which happen to point almost toward the vertex, as well as spurious tracks created by background coincidence. The candidate is vetoed if it is missed in two or more consecutive emulsions, i.e., if it would be in the CCD field of view and also well-separated from nearby tracks, but is not measured. Thus, (g) is accepted but (h) is rejected. The detection efficiency for well-separated tracks is 99% on average, and tracks are measured in no more than 25 separate emulsions, so the probability of two or more consecutive misses in a real track is typically less than  $25 \times (0.01)^2 \sim 0.25\%$ .

**Close Pair** A background measurement in proximity to a real track can imitate a pair of tracks. To be accepted as a real track, a cluster must have at least two measurements which belong uniquely to that cluster. In Fig. 15, clusters (c) and (d) pass this test, but (f) does not. An exception is made for tracks with a single unique measurement if the unique measurement occurs in the most downstream measurable emulsion. Such a track is likely to be one of a close pair which is resolved just before it leaves the field of view.

Systematic errors are estimated on a track by track basis using background and efficiency information derived from the reconstruction. Values are computed for four kinds of systematics:

**Missing Measurements** There is a probability of about 1% that a track will be missed in a particular emulsion. If the track can only be measured in two



emulsions and is missed in one, it is then measured only once and is incorrectly considered to be background. If the track can be measured in three emulsions, according to the consistency tests above, it is only rejected if it is missed in two consecutive emulsions. The probability that a track will be missed because of missing measurements drops rapidly with the number of possible measurements, and the expected number of missed tracks is very nearly twice the number of double measurements times the measurement inefficiency. For Event 20-06, the expected number of missed tracks is approximately  $2 \times 171 \times 0.06 = 2.0$  tracks, compared to 762 measured tracks in all.

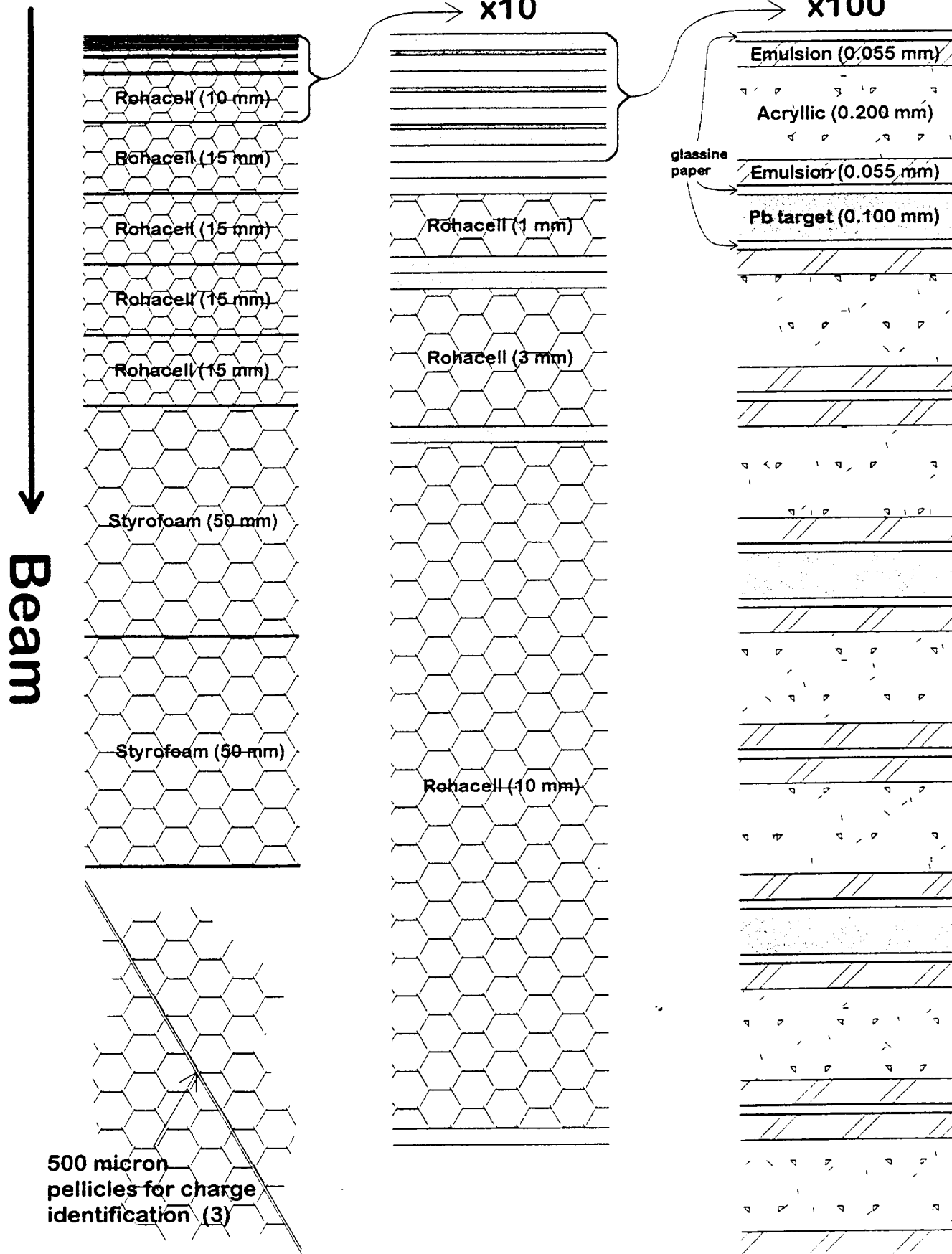
**Spurious Doubles** Spurious doubles are created when a background measurement in one emulsion coincides with a background measurement in another, or when a random track which nearly points back to the vertex passes through the field of view. Accidental coincidences are cut if the supposed track is missing in at least two consecutive emulsions in which it could be measured if it were real. This cut is only about 50% efficient, since it can only be used if there are four or more measurable emulsions. In Event 20-06, 8 spurious doubles were detected, suggesting another 8 remain undetected.

**False Close Pairs** The main source of false close pairs appears to be single measurements incorrectly identified in image processing as two measurements. False close pairs are cut if one of the tracks has a missing downstream measurement. About 6% (48) tracks were cut on this basis. False close pairs remain undetected in that sample of tracks which branch only in the most downstream emulsion. Event 20-06 has 118 such tracks. We estimate that  $\frac{1}{2} \times 118 \times 0.062 = 3.6$  false close pairs remain undetected.

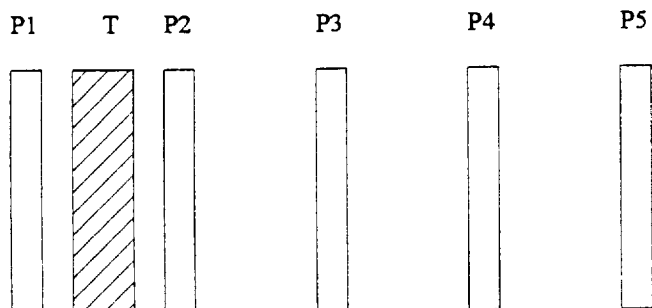
**Unresolved Close Pairs** These are tracks which are too close together to be

optically resolved. Using the measured instrumental pair resolution, we estimate that we undercount the number of pion tracks by roughly 6 in a typical high-multiplicity event. In addition, we expect that 10-20 close election pairs from  $\pi^0$  decay are counted as single tracks.

# EMU-13 Emulsion Chamber



Beam 



Scanning

Tracing

Acquisition (Low Power)

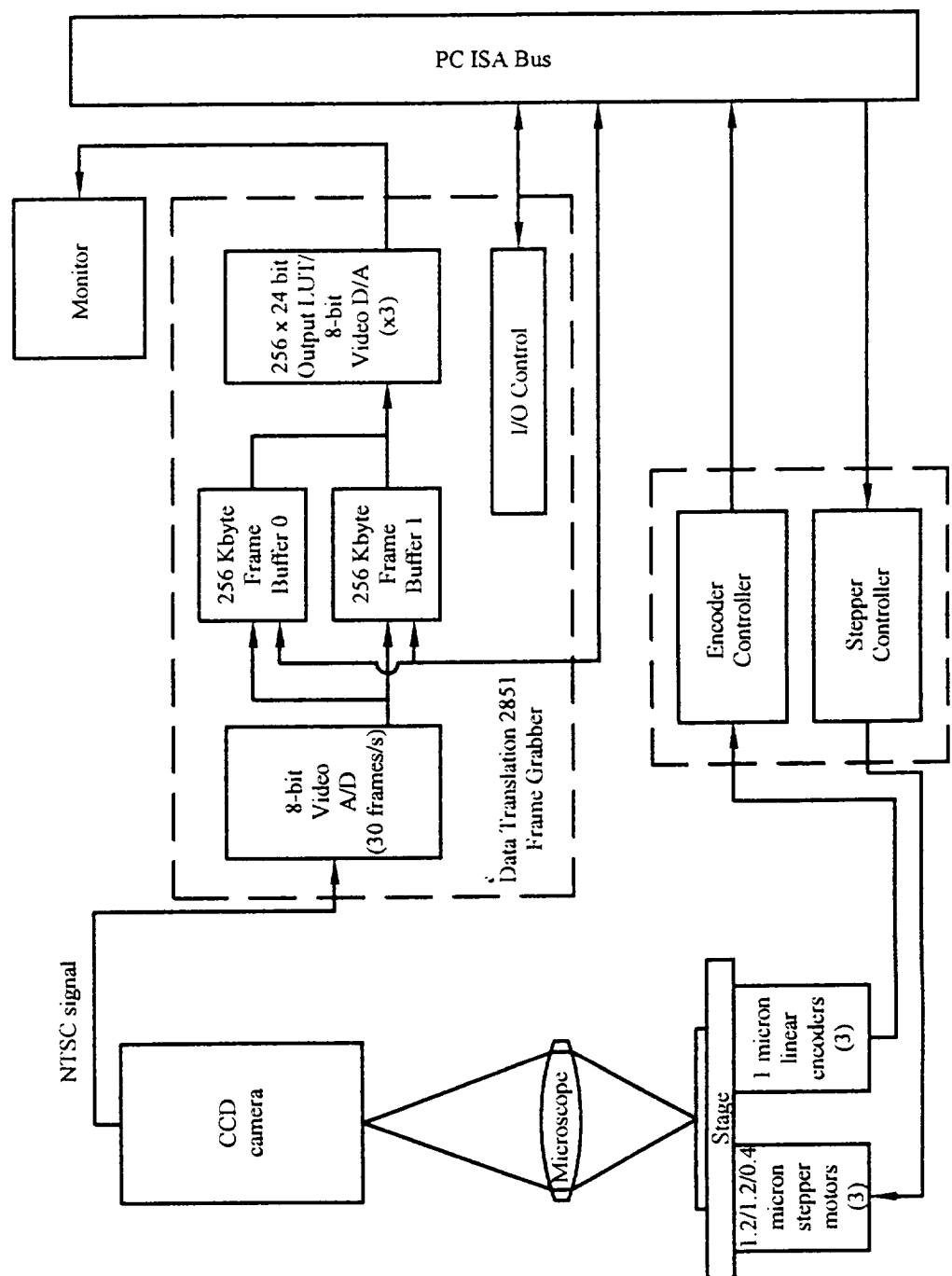
Acquisition (High Power)

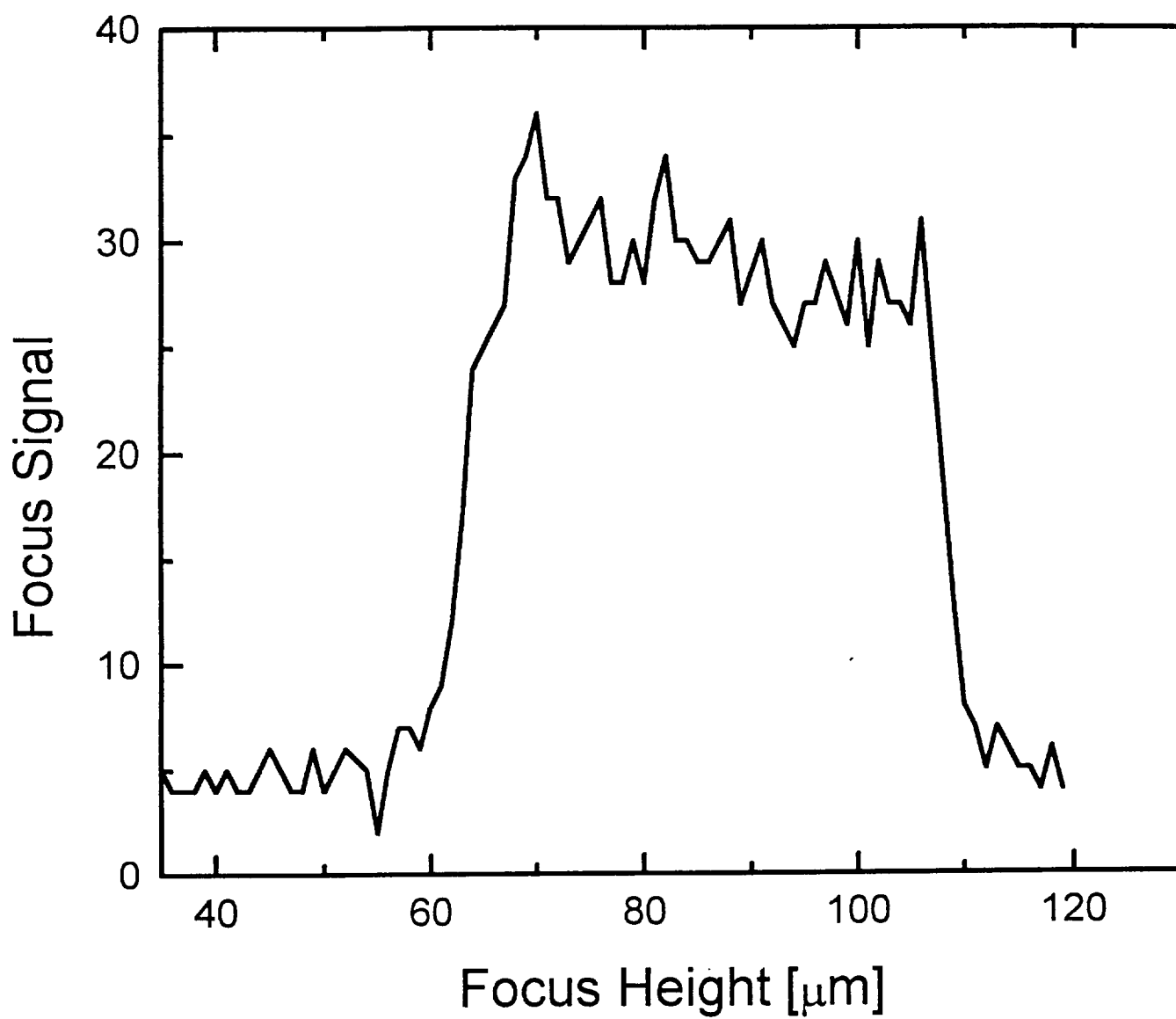
Vertex Finding

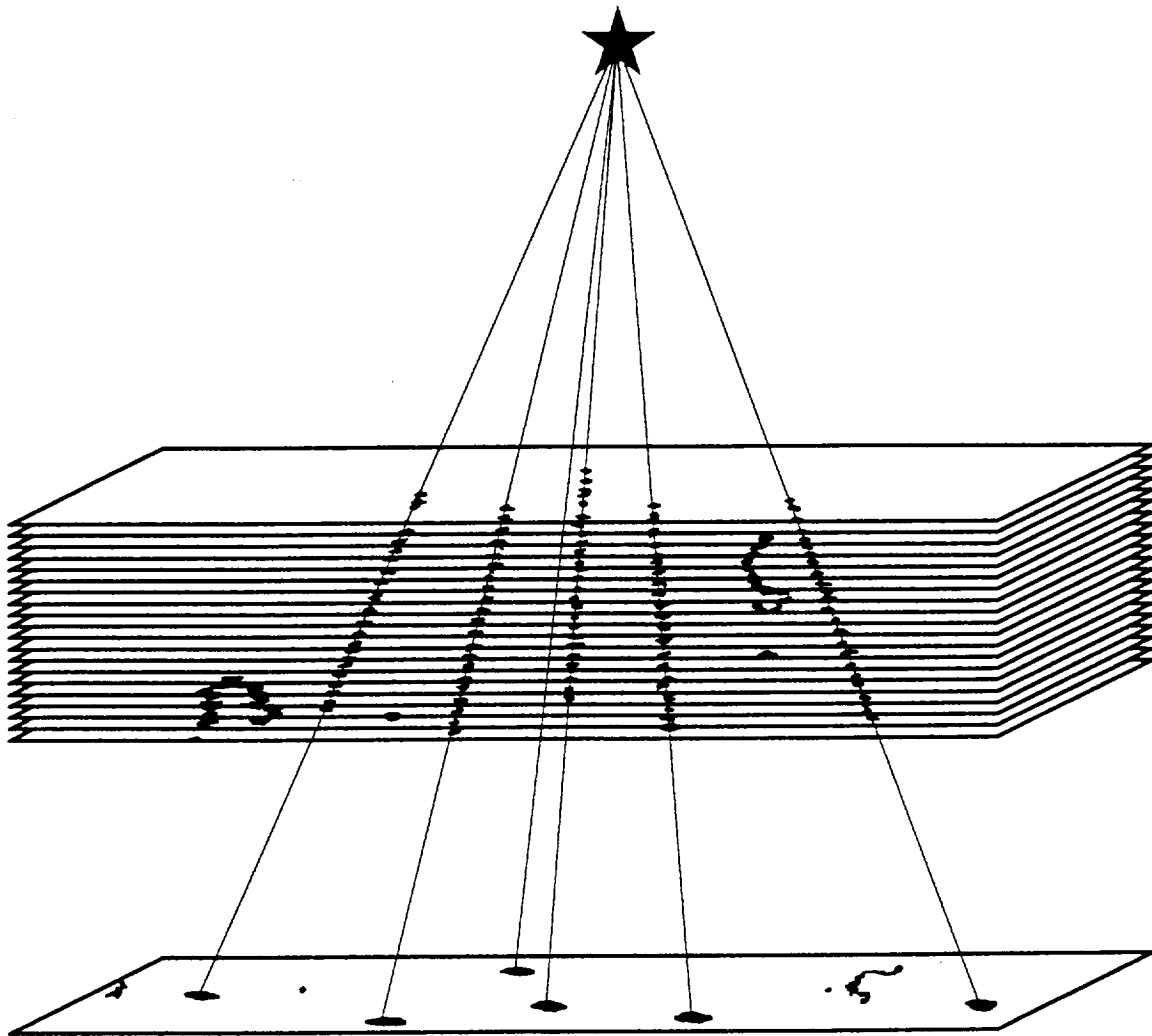
Stage 2 Image Processing

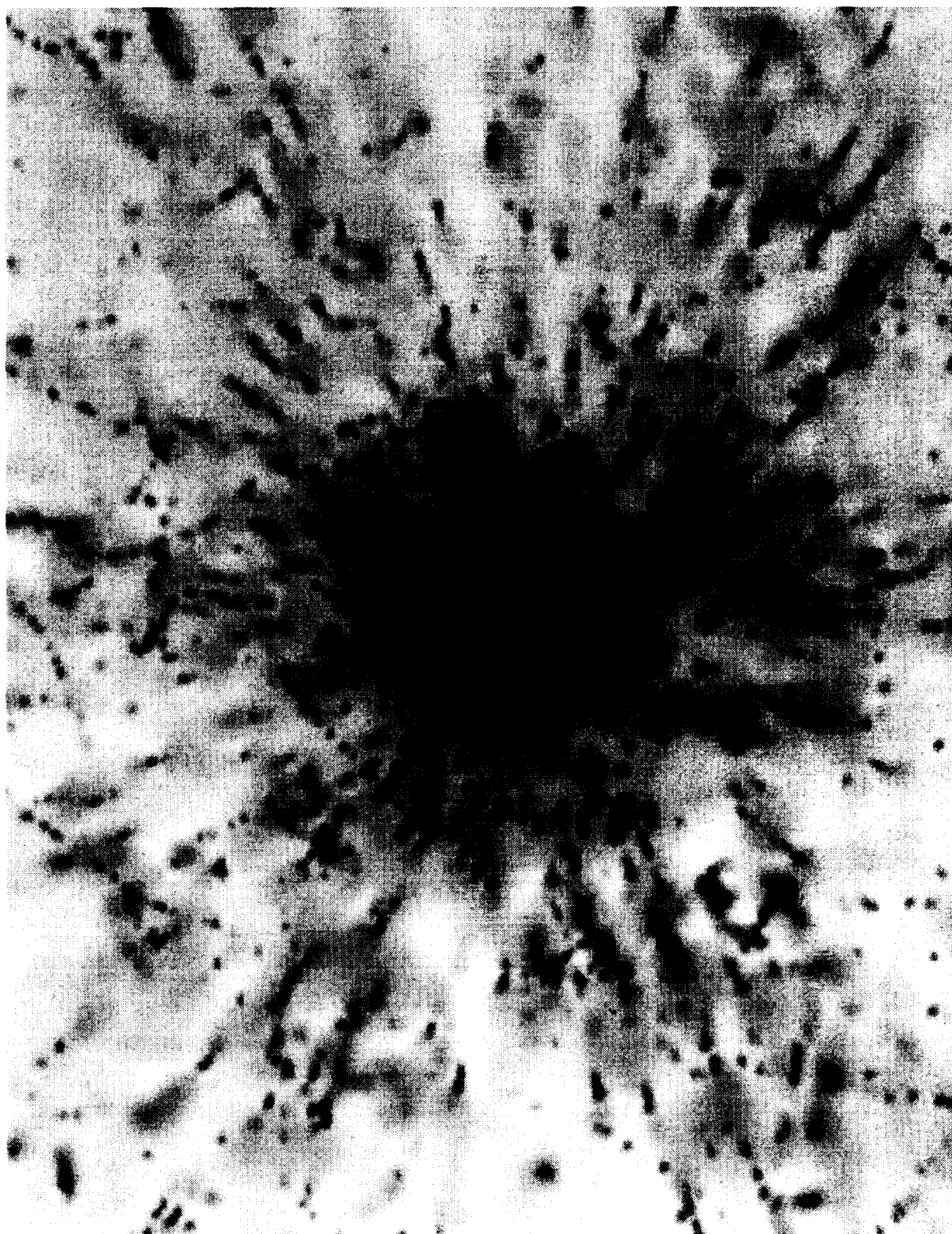
Plate Fitting

Reconstruction

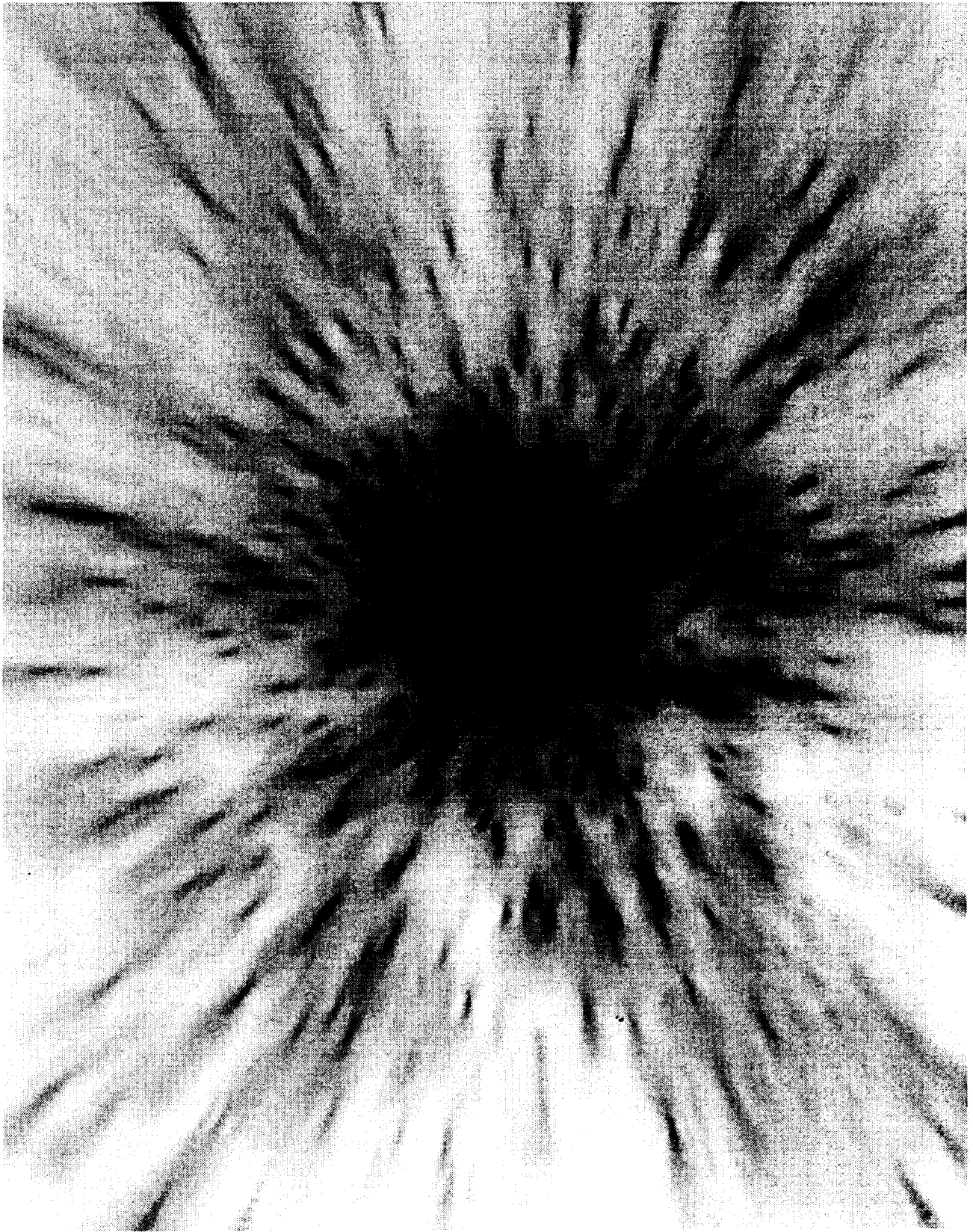


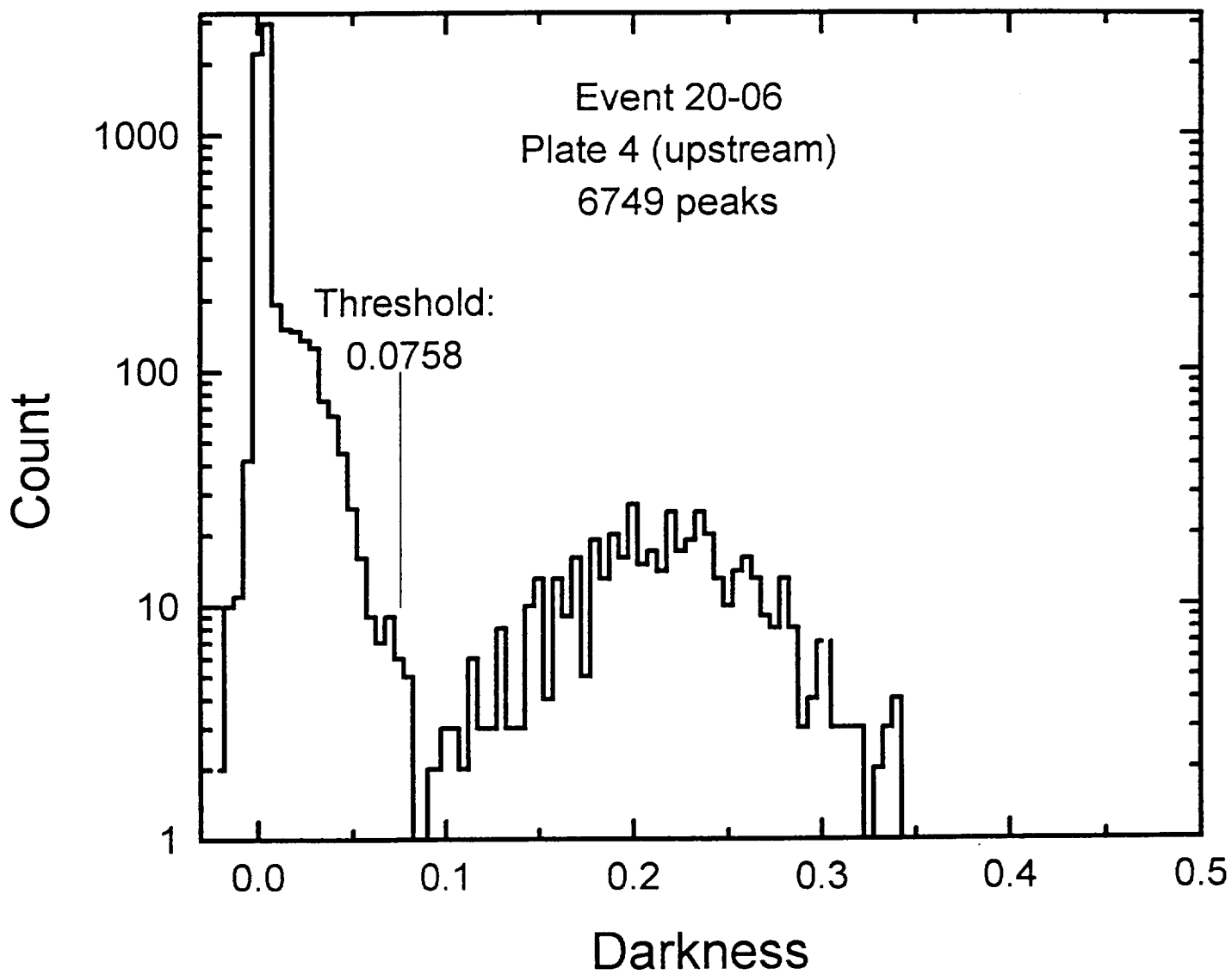












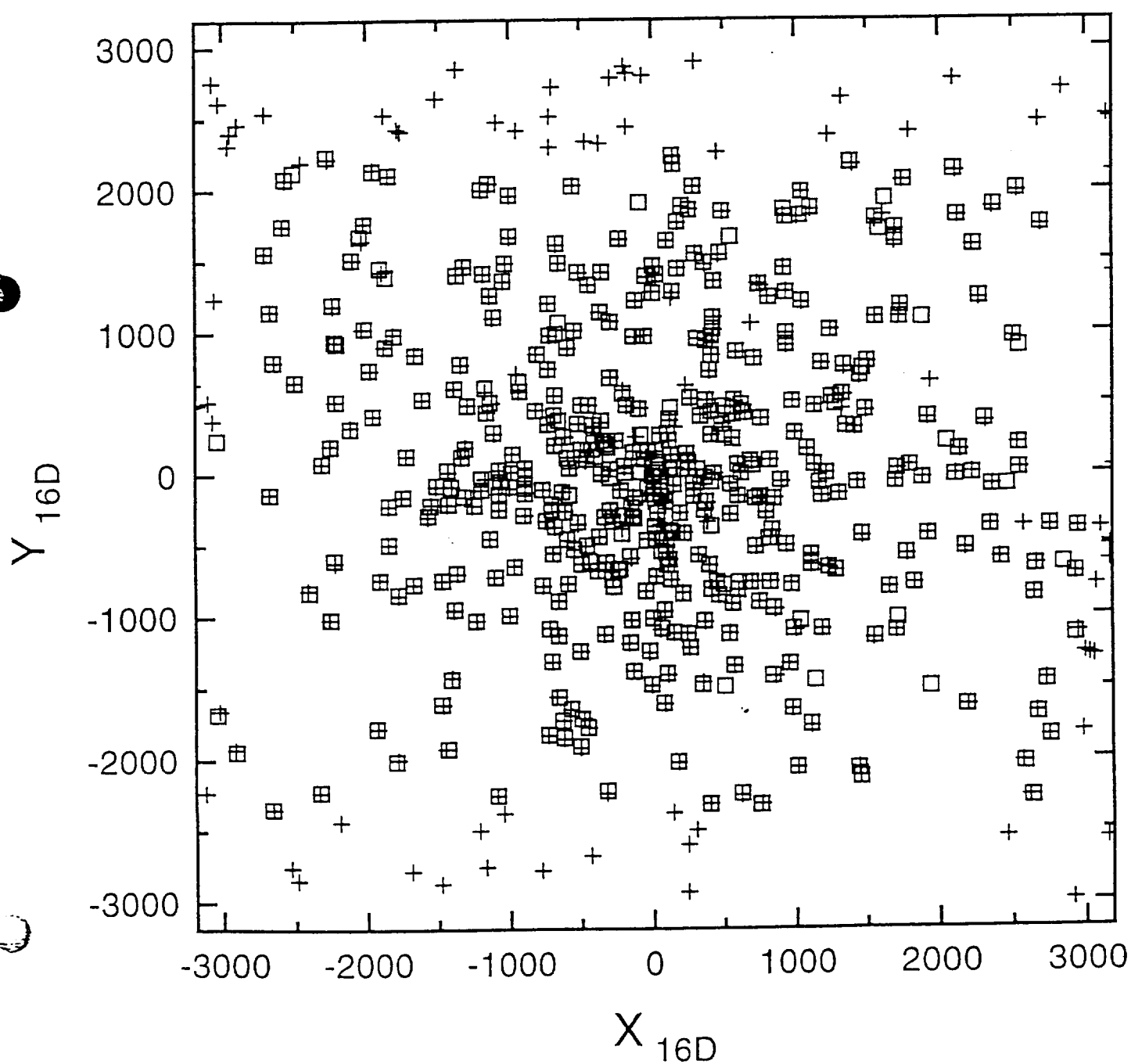
## Event 2006

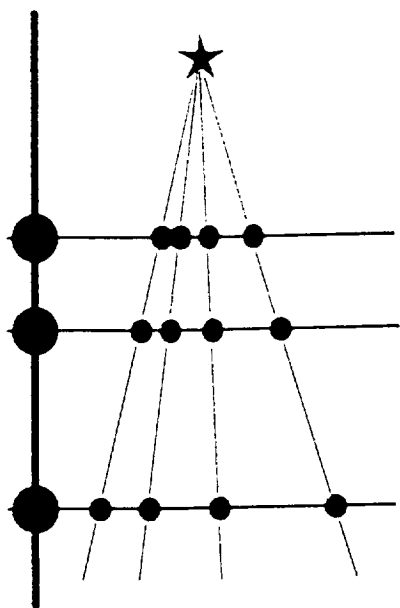
 $\square$  5U

+ 4U

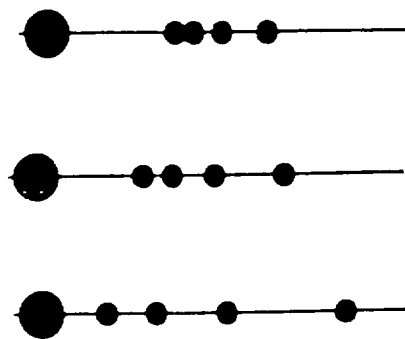
 $\Delta x = -0.97$  $\Delta y = 0.41$ 

scale = 1.2672

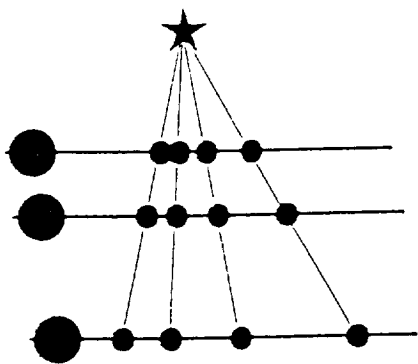




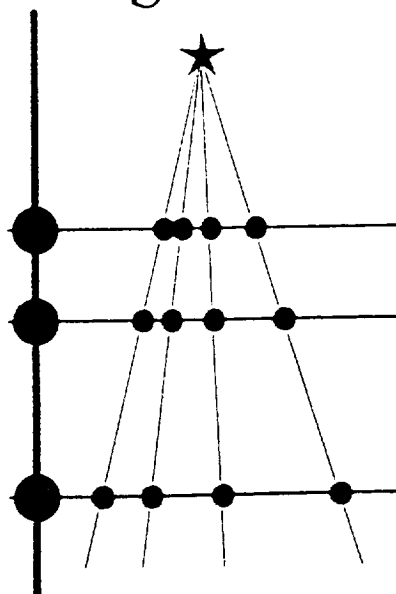
a



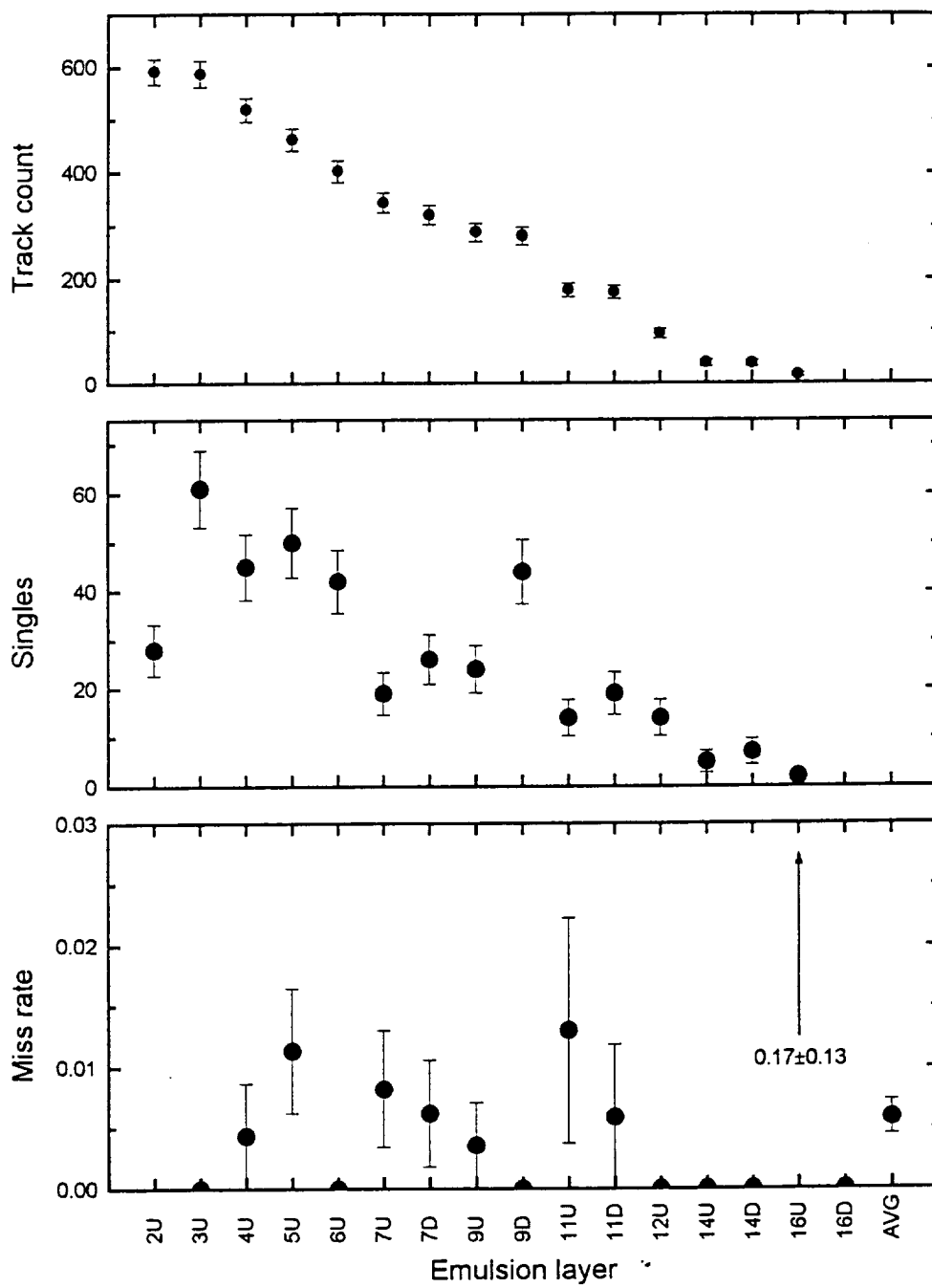
b

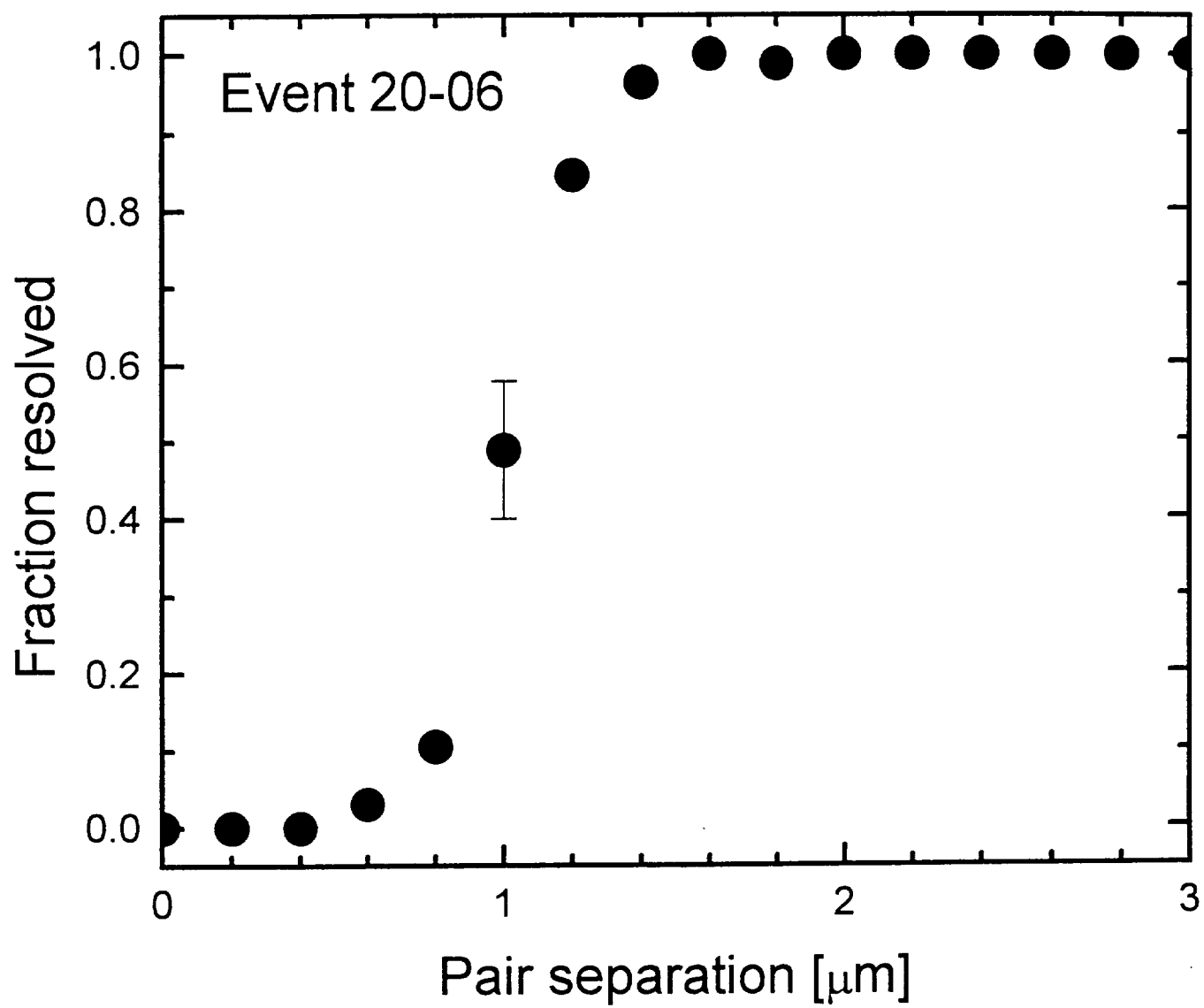


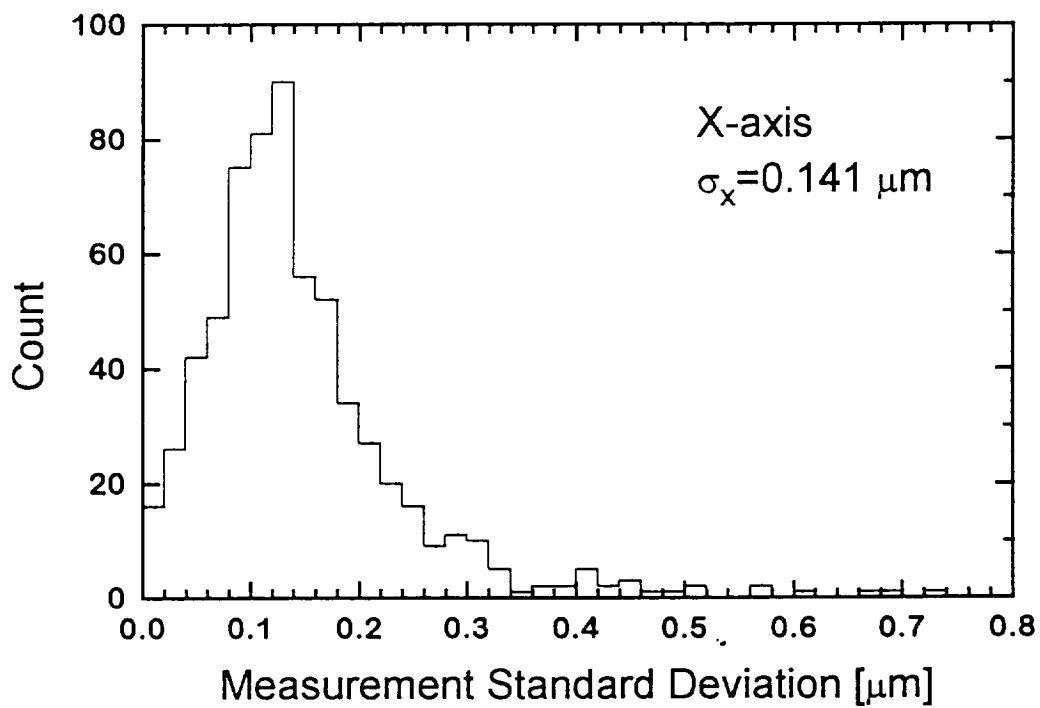
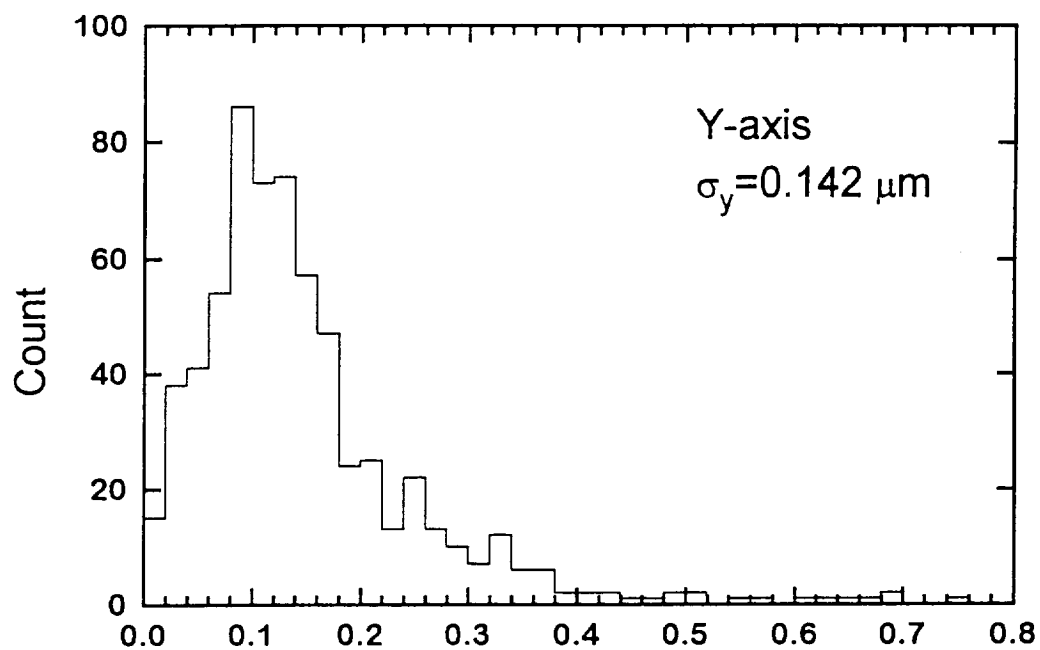
c

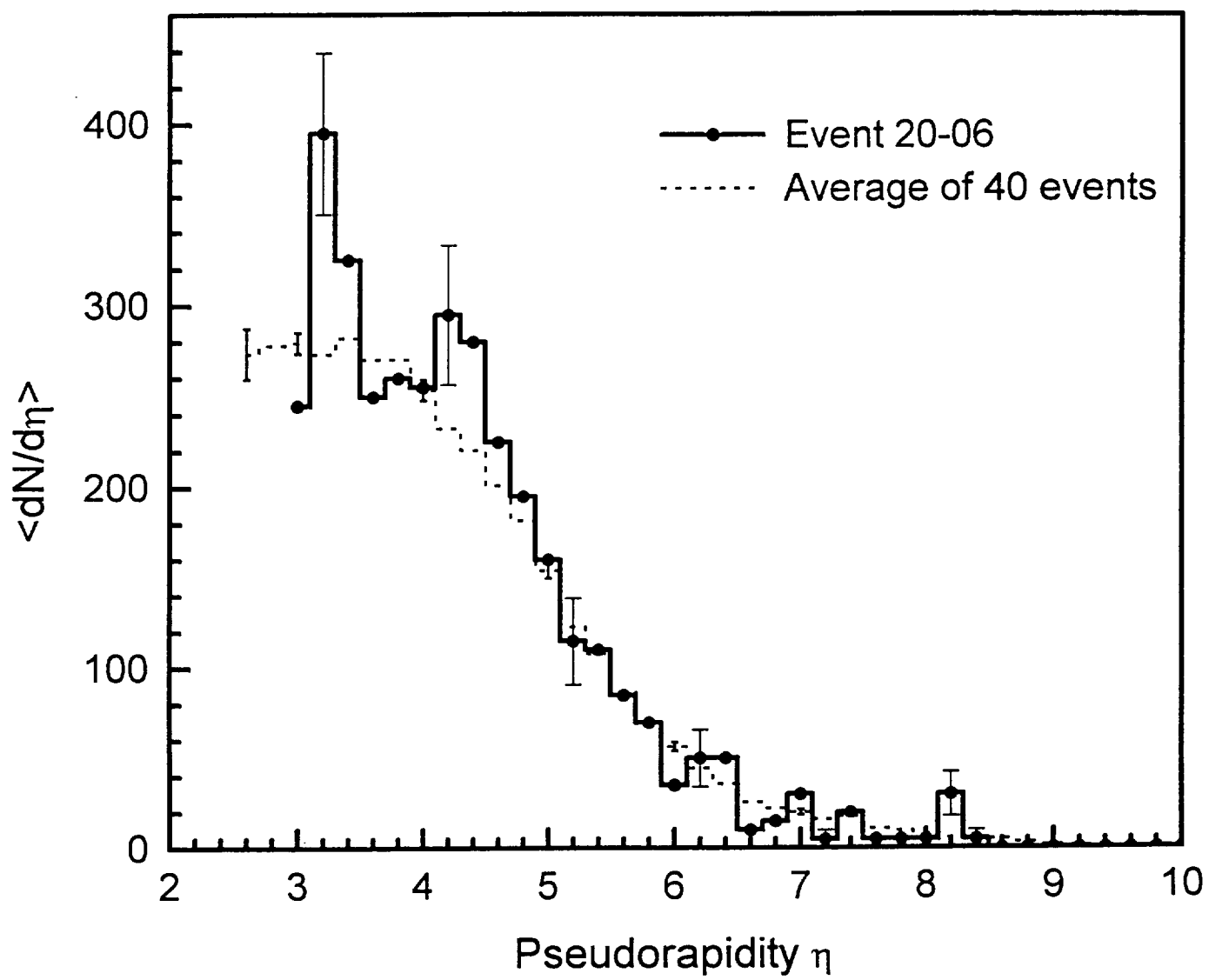


d

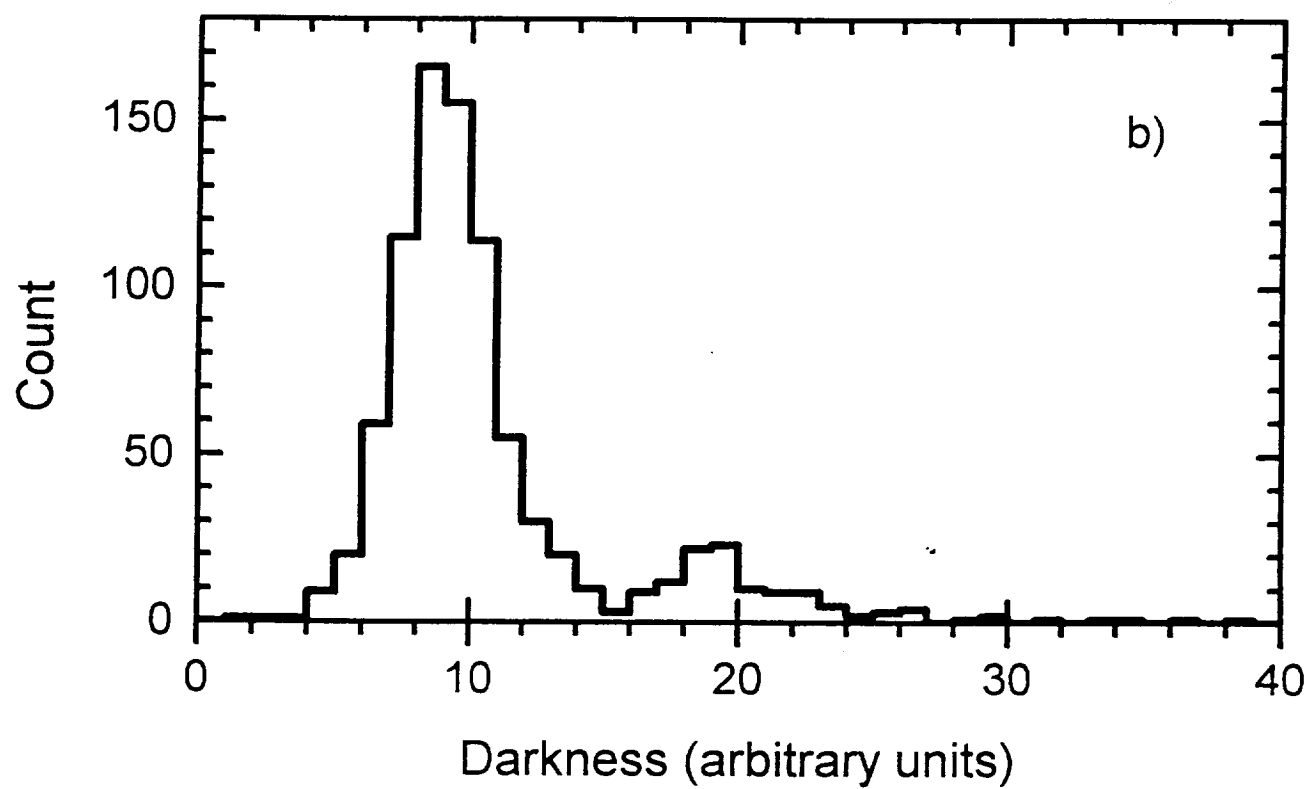
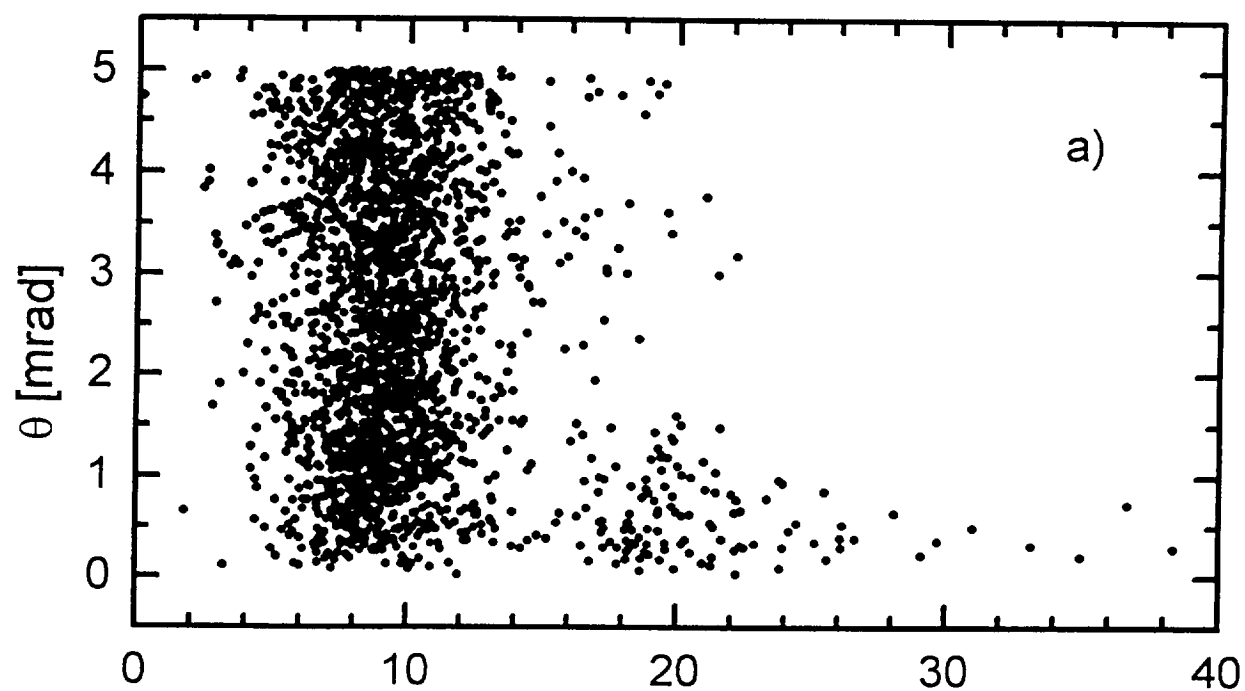


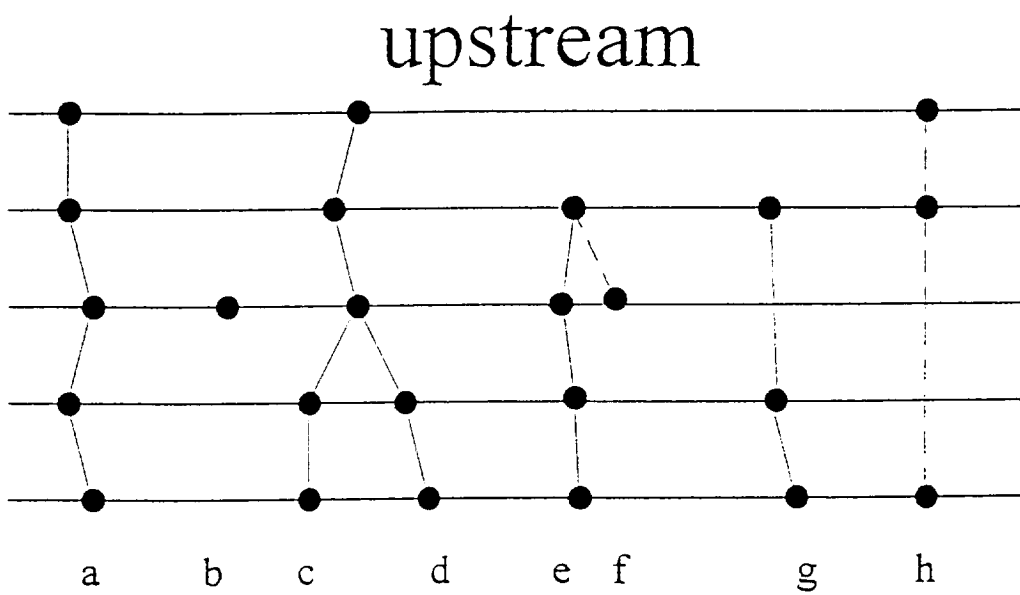












downstream