SLAC–PUB–7790
April 1998

Track Fitting and Antennas *

Richard Blankenbecler

Stanford Linear Accelerator Center
Stanford University, Stanford, California 94309
rzbth@slac.stanford.edu

Abstract

The problem of fitting track data is transformed into a problem in antenna theory. This latter well-studied problem is characterized as an antenna array that is receiving a narrow band signal from multiple distant sources. The goal here is to count the number of sources and determine the angle of each source relative to the array. Similarly, the original problem of track fitting is to count the number and to determine the location and angle of each track, but in the presence of noise and finite detection efficiency. However, an additional complication in fitting tracks is that in a magnetic field, the radius of curvature of the track must also be determined. This is shown to map into an extended antenna problem of analyzing 'chirped' or frequency modulated sources. A somewhat detailed development and discussion of track parameter estimation is then given.

*Work supported by the Department of Energy, contract DE–AC03–76SF00515.

1 Introduction and Motivation

In the real world, the measurement of global characteristics of an image over a large volume or area are beset by a number of difficulties. In the paper by Aghajan and Kailath[1] an elegant and useful method for fitting multiple lines in a two-dimensional image was given that exploits the analogy to the problem of an antenna array that is receiving a narrow band signal from multiple distant sources. This work was an extension of the work by Roy and Kailath[2] on the "ESPRIT" method (Estimation of Signal Parameters via Rotational Invariance Technique). Other work in this field has been done by Lou, Hassebrook, Lhamon, and Li[3] and Kumaresan and Tufts[4].

A simple introduction to this area of research is given in Appendix A and Appendix B. These appendices do not discuss the more sophisticated methods developed in the above papers; they are meant to indicate both the logical connection to the method proposed in the present paper as well as contrast the methods used.

In many high energy physics experiments, it is necessary to measure the characteristics of tracks produced by particles as they transit a detector. These 'images' are corrupted by noise and by the finite detection efficiency of the active elements of the detector. Track fitting commonly proceeds by two stages: first estimating the number of tracks (lines) and their parameters in an event, and then passing this information to a more elaborate fitting procedure to extract accurate values for the parameters of each track.

An example problem is illustrated in Figure 1. Each square 'hit' denotes a response from one of the elements in the detector volume. It is quite easy to see (or at least imagine) that there are two tracks in this event. It is also evident that the detector has both noise and a finite detection probability. The mathematical problem is to develop an algorithm that will count the number of tracks and fit the shape of the tracks while ignoring the noise to the maximum extent possible.

A standard approach to the first stage estimate of the number of tracks and their parameters is the Hough transform, described and extended in the paper by Pao, Li, and Jayakumar[5]. In this paper, an improvement is given that extracts the same information using a more efficient algorithm than the standard Hough transform.

The second stage can be handled by a variety of methods too numerous and complicated to mention. The development of one method, called Deformable Templates, or Elastic Arms, can be found in M. Ohlsson, C. Peterson, and A. Yuille[6]. Extensions of this method have been given by M. Ohlsson[7] and R. Blankenbecler[8].

In this paper, the problem of fitting track data is transformed into an analogous problem in antenna signal analysis in which the goal is to count the number of radiating sources and determine the angle of each source relative to an antenna array by suitable manipulation of the received signal. Similarly, the basic problem of track fitting is to count the number and to determine the location and angle of each track in the presence of noise and finite efficiency. These two different problems are illustrated in Figure 2. The upper diagram schematically defines the antenna problem, while the lower diagram illustrates the simulated antenna used in our treatment of track fitting.

For curved tracks, an extension of the above concepts must be developed. It will be shown that in this case, the analogous signals incident on the antennas are frequency modulated. The detection of a restricted class of 'chirped' signals is discussed in a form useful for the present problem by E. T. Jaynes[10], G. L. Bretthorst[11] and Erickson, Neudorfer and Smith[12].

Standard images are two dimensional and the discussion here will explicitly treat only this case. However detectors measure three dimensional tracks. Thus the present analysis deals separately with the two transverse projections of such data. The treatment of the full three dimensional case together with the additional information and constraints that one projection imposes on the other will be given later.

2 Single Curving Track with Noise

Consider a two dimensional image plane of area $Y * Z$ consisting of pixels that can take values of '1' and '0'. These values are given by the matrix $I_{j,m}$, where $0 < j < J$ sweeps out the y-direction, and $0 < m < M$, the z-direction. As advertised, $I_{j,m} = 1$ or 0. First we will discuss an image consisting of a smooth line, or track, together with noise pixels, or 'outliers', for which $I_{j,m} = 1$. Thus

$$I_{j,m} = (1 - \mathbf{e}_{j,m}) \qquad \text{for} j = t(m) \tag{1}$$

$$I_{j,m} = \mathbf{n}_{j,m} \qquad \text{otherwise. .} \tag{2}$$

where the equation of the track is $j = t(m)$. In the simple linear case, $j = t_0 + t_1 m$, with t_1 measuring the y-z slope. Our treatment will hold for a general track shape. The fluctuating variable $\mathbf{e}_{j,m}$ measures the detection *inefficiency*. Thus $\mathbf{e}_{j,m} = 0$ if the pixel 'fired', and $\mathbf{e}_{j,m} = 1$ if it did not. Similarly, the noise is given by the fluctuating variable $\mathbf{n}_{j,m}$. Since there is no apriori way of identifying noise hits with track hits, any procedure must treat all data points the same. It is the analysis itself that must fit the real hits and ignore the noise hits to the maximum extent possible.

Now formally define a pair of 'signals' at the mth row which is given by the sum over all the nonzero pixels in the y-direction

$$c_m = \sum_j I_{j,m} \cos[\delta\phi\, j] \tag{3}$$

$$= (1 - \mathbf{e}_{t(m),m}) \cos[\delta\phi\, t(m)] + \sum_{j \neq t(m)} \mathbf{n}_{j,m} \cos[\delta\phi\, j] , \tag{4}$$

and

$$s_m = \sum_j I_{j,m} \sin[\delta\phi\, j] \tag{5}$$

$$= (1 - \mathbf{e}_{t(m),m}) \sin[\delta\phi\, t(m)] + \sum_{j \neq t(m)} \mathbf{n}_{j,m} \sin[\delta\phi\, j] , \tag{6}$$

where $\delta\phi$ is a fixed parameter to be chosen later. The difference between the above signals and the signals in a perfect detector with no noise is

$$\begin{aligned} \delta c_m \quad &\equiv \quad c_m \; - \; \cos[\delta\phi\, t(m)] && (7) \\ &= \quad -\mathbf{e}_{t(m),m} \cos[\delta\phi\, t(m)] \; + \sum_{j \neq t(m)} \mathbf{n}_{j,m} \cos[\delta\phi\, j] \; , && (8) \\ \text{and} \qquad \delta s_m \quad &\equiv \quad s_m \; - \; \sin[\delta\phi\, t(m)] && (9) \\ &= \quad -\mathbf{e}_{t(m),m} \sin[\delta\phi\, t(m)] \; + \sum_{j \neq t(m)} \mathbf{n}_{j,m} \sin[\delta\phi\, j] \; , && (10) \\ \text{and define} \qquad (\delta o_m)^2 \quad &\equiv \quad (\delta c_m)^2 + (\delta s_m)^2 \; . && (11) \end{aligned}$$

First preform the statistical average of the inefficiency and noise variables which range between zero and one. Then average over all possible parameters of the track which drives the cross terms to zero. The final ensemble average is

$$\sigma^2 \; = < (\delta o_m)^2 > = < \mathbf{e}_m^2 > \; + \, (J-1) < \mathbf{n}_m^2 > \; , \qquad (12)$$

where σ is the measure of the expected fluctuation of the quantity δo_m . The first term is the expected fluctuation coming from the inefficiency of detecting the track, while the second term measures the expected noise from the remaining $(J-1)$ pixels.

Thus given the true equation of the track $t(m)$, the probability that the data set $O = \{o_m\}$ will occur is just the probability that the fluctuations will make up the difference:

$$p(O|t,\sigma) \; = \; \prod_{m=0}^{M-1} \frac{1}{\sqrt{2\pi\ \sigma^2}} \exp[-\frac{1}{2\sigma^2}(\delta o_m)^2] \; , \qquad (13)$$

where the argument t stands for all the parameters describing the track.

Conversely, given the noise level σ and the data O , the joint likelihood of the parameters of the equation of the track, $t(m)$, is

$$L(t,\sigma) \; \propto \; \exp[-\frac{1}{2\sigma^2} \sum_{m=0}^{M-1} (\delta o_m)^2] \; . \qquad (14)$$

2.1 Approximations

In order to extract the general behavior of this result, certain approximations will be made at this juncture. Expand the quadratic expression δo_m^2 . The first term, o_m^2 , depends only on the data, not on the parameters to be fitted. Except for particular values of the parameters of the track, the last term can be approximated by[13]

$$\sum_0^{M-1} \cos^2[\delta\phi\, t(m)] \sim \sum_0^{M-1} \sin^2[\delta\phi\, t(m)] \sim \frac{1}{2}\, M \ . \tag{15}$$

The important dependence of the likelihood function on the track parameters then arises only from the cross terms and

$$L(t,\sigma) \ \propto \ \exp\left[\frac{1}{\sigma^2}\sum_0^{M-1}(c_m \ \cos[\delta\phi\, t(m)] + s_m \ \sin[\delta\phi\, t(m)]\right] \ . \tag{16}$$

First we concentrate our interest on the parameter $t(0)$, the intercept of the track. Define $\theta = \delta\phi\, t(0)$ so that $\delta\phi\, t(m) = \delta\phi\,(t(m) - t(0)) + \theta = \delta\phi\, \Delta t(m) + \theta$. The likelihood function L becomes

$$L(t,\sigma) \ \propto \ \exp\left[\frac{1}{\sigma^2}\ \Sigma\ \right], \tag{17}$$

$$\text{where} \qquad \Sigma = \sum_0^{M-1}(c_m \cos[\delta\phi\, \Delta t(m) + \theta] + s_m \sin[\delta\phi\, \Delta t(m) + \theta] \tag{18}$$

$$\Sigma = \Sigma_c \ \cos\theta - \Sigma_s \ \sin\theta \ , \tag{19}$$

with

$$\Sigma_c = \sum_0^{M-1}(c_m \cos[\delta\phi\, \Delta t(m)] + s_m \sin[\delta\phi\, \Delta t(m)]) \tag{20}$$

$$\Sigma_s = \sum_0^{M-1}(c_m \sin[\delta\phi\, \Delta t(m)] + s_m \cos[\delta\phi\, \Delta t(m)]) \ . \tag{21}$$

The value of $t(0)$ (determined to within a branch ambiguity of the arctangent) that maximizes the likelihood function $L(t,\sigma)$, and the corresponding maximum of Σ , are

$$\tan\theta = -\Sigma_s/\Sigma_c \quad \text{and} \quad \Sigma_{max} = \sqrt{\Sigma_c^2 + \ \Sigma_s^2} \ . \tag{22}$$

In solving for θ , the branch of the tangent function must be chosen so that the second derivative of Σ is positive. Also note that this latter quantity can be rewritten as

$$
\begin{aligned}
\Sigma^2_{max} &\equiv MC(t) & (23)\\
&= \frac{1}{M}\sum_{m,n}(c_m c_n + s_m s_n)\cos[\delta\phi\,(\Delta t(m) - \Delta t(n))] \\
&\qquad\qquad +(c_m s_n - s_m c_n)\sin[\delta\phi\,(\Delta t(m) - \Delta t(n))]) \ . & (24)
\end{aligned}
$$

On the other hand, one could drop immediate interest in the parameter $t(0)$. This "nuisance" parameter should then be integrated out of the likelihood function. To that end introduce the 'reduced' likelihood function L_0 and carry out the integral to yield

$$
L_0(t,\sigma) \ \propto \ \int_0^{2\pi}\frac{d\theta}{2\pi}\exp\left[\frac{1}{\sigma^2}\ \Sigma\ \right], \qquad (25)
$$

$$
\propto \ I_0[\sqrt{MC(t)}/\sigma^2] \ , \qquad (26)
$$

where $I_0(x)$ is a standard Bessel function and $C(t)$ is the same function introduced in eqn(24). The function I_0 is a monotonically increasing function of its argument. The likelihood functions L and L_0 are different in form, because the questions asked were different, but in both cases the optimum values of the remaining track parameters, t_i $(i \neq 0)$, are determined by the maximum of the same function, namely $C(t)$.

The function $C(t)$ is a generalization of the Schuster[9] *periodogram* used in spectral analysis of time series. E. T. Jaynes[10] has applied this function to the analysis of frequency modulated signals and termed it a *chirpogram*. In the present case, the name *trackogram* seems to be descriptive. Note that all of our general discussion holds for any track function. For a curving (quadratic) track, $t(m) = t_0 + t_1 m + t_2 m^2$, and

$$
t(m) - t(n)) \ = \ \Delta t(m) - \Delta t(n)) \ = \ [\delta\phi\, t_1(m-n) + \delta\phi\, t_2(m^2-n^2)] \ . \qquad (27)
$$

Hence the likelihood function is a maximum for the values of the track parameters t_1 and t_2 which maximize $C(t)$.

Note also that the calculation of $C(t)$ from the double sum in eqn (24) requires $\sim M^2$ steps, whereas its evaluation from Σ_c and Σ_s as in eqn (21) requires only $\sim 2\,M$ steps, a considerable savings for large M . Finally, the maximum value of $C(t)$ can be estimated to be $\sim M$.

3 Multiple Curving Tracks with Noise

Now consider the case of D tracks that are described by the functions

$$t_d(m)\ , \qquad \text{for } 0 \le d < D\ . \tag{28}$$

The 'signal' at the mth row is now given by the sum over all nonzero pixels on that row:

$$\begin{aligned} c_m^d &= \sum_d (1-\mathbf{e}_{t_d(m),m})\cos[\delta\phi\, t_d(m)] \ + \ \sum_{\neq} \mathbf{n}_{j,m}\cos[\delta\phi\, j]\ , && (29)\\ \text{and} \qquad s_m^d &= \sum_d (1-\mathbf{e}_{t_d(m),m})\sin[\delta\phi\, t_d(m)] \ + \ \sum_{\neq} \mathbf{n}_{j,m}\sin[\delta\phi\, j]\ , && (30)\end{aligned}$$

where $\sum_{\neq}$ means that all terms for which $j = t_d(m)$, $0 \le d < D$, are omitted. The differences between the above signal and the signal in a perfect detector with no noise for the these two signals are

$$\delta c_m \ = \ c_m \ - \ \sum_d \cos[\delta\phi\, t_d(m)]\ , \qquad \delta s_m \ = \ s_m \ - \ \sum_d \sin[\delta\phi\, t_d(m)]\ , \tag{31}$$

and the expected total fluctuation is

$$\sigma^2 \ = \ <(\delta o_m)^2> \ = \ D<\mathbf{e}_m^2> \ + (J-D)<\mathbf{n}_m^2>\ , \tag{32}$$

which depends both upon D and J , as expected. The likelihood function is again of the form

$$L(\vec{t},\sigma) \ \propto \ \exp[-\frac{1}{2\sigma^2}\sum_0^{M-1}(\delta o_m)^2]\ , \tag{33}$$

where the argument $\vec{t}$ stands for all the parameters describing each of the D tracks.

3.1 Approximations

Following the previous line of argument and approximations, the square of the track terms involves

$$\sum_{d1,d2} \sum_{0}^{M-1} \cos[\delta\phi\, t_{d1}(m)] \cos[\delta\phi\, t_{d2}(m)] \sim \frac{1}{2}\, M\ D\ , \tag{34}$$

with a similar result holding for the sin terms. The nondiagonal terms average to zero. The dependence of the likelihood function on the track parameters then again arises primarily from the cross term between the data and the track term. This cross term then factorizes:

$$L(\vec{t},\sigma) = \prod_d L(t_d,\sigma) = \prod_d \exp[\Sigma(d)/\sigma^2] \tag{35}$$

$$\text{where} \quad \Sigma(d) = \sum_m (c_m\ \cos[\delta\phi\, t_d(m)] + s_m\ \sin[\delta\phi\, t_d(m)])\ . \tag{36}$$

At this point the previous discussion can be followed in detail and the results simply copied over. Again define $\theta_d = \delta\phi\, t_d(0)$ and $\Delta t_d(m) = t_d(m) - t_d(0)$ so that

$$\Sigma(d) = \Sigma_c(d)\ \cos\theta_d - \Sigma_s(d)\ \sin\theta_d\ , \tag{37}$$

where $\Sigma_c(d)$ and $\Sigma_s(d)$ are given by eqn(21) but with the track function replaced by $t_d(m)$. The values of $t_d(0)$ that maximize Σ , and the resultant Σ_{max} are

$$\tan[\delta\phi\, t_d(0)] = -\Sigma_s(d)/\Sigma_c(d) \quad \text{and} \quad \Sigma_{max} = \sum_d \Sigma(d)\ . \tag{38}$$

where $\Sigma(d)^2 = \Sigma_c(d)^2 + \Sigma_s(d)^2 = MC(t_d)$ with $C(t_d)$ defined in eqn(24).

If the $t_d(0)$ are treated as nuisance parameters, then one has

$$L_0(\vec{t},\sigma) \propto \prod_d \int_0^{2\pi} \frac{d\theta_d}{2\pi}\ \exp\left[\frac{1}{\sigma^2}\sum_m \Sigma(d)\right] \tag{39}$$

$$\propto \prod_d I_0[\sqrt{MC(t_d)}/\sigma^2]\ . \tag{40}$$

The reduced likelihood function has factored into a product of independent distributions.

4 Numerics and Examples

It is convenient to scale the parameters so that the two dimensional plane containing the image has unit dimensions. To that end define $z = m/(M-1)$ and

$$\delta\phi\, t(m) \;=\; \delta\Phi\, y(z), \qquad \text{where} \qquad \delta\Phi = \delta\phi\, J \tag{41}$$

$$\text{with} \qquad 0 \;<\; y(z) < 1, \qquad \text{and} \qquad 0 < z < 1. \tag{42}$$

All the hits now lie in the unit square. In order to determine the best estimate of the track parameters, the function $C(t)$ must be studied and its maximum value determined. There are several approaches to this problem. We have chosen to use a simple histogramming technique coupled with the Simplex method since they directly generalize to more complicated track forms. The Simplex method is discussed in the book Numerical Recipes[14].

Since the function to be maximized, $C(t)$, does not depend upon the intercept of the track, it is expeditious to change the parameterization of the track so that the two degrees of freedom are as independent as possible. The midpoint slope of the track and the curvature around this value are suitable fitting parameters. The parameterization of the track is therefore changed to

$$y(z) \;=\; k_0 + \frac{1}{2}k_1 z(1+z) + \frac{1}{2}k_2 z(1-z)\;, \qquad \text{with} \tag{43}$$

$$t_0 \;=\; Jk_0, \quad (M-1)t_1 = \frac{J}{2}(k_1+k_2), \quad (M-1)^2 t_2 = \frac{J}{2}(k_1-k_2)\;, \tag{44}$$

thus $C(t)$ becomes $C(k)$. The slope parameter at the midpoint in z is k_1; the boundary conditions on the track are $y(0) = k_0$, and $y(1) = k_0 + k_1$. The track fitting parameters are conveniently chosen to be k_1 and k_2. The parameter $\delta\Phi$ is arbitrary, chosen during the fitting process. This will be discussed further below. The first step is to assume a value for $\delta\Phi$ and compute the vectors c_m and s_m from the data.

The net transverse displacement in crossing the detector is $y(1) - y(0) = k_1$. It will be shown that k_1 is well determined by the study of $C(k_1, k_2)$ while, as has been previously noted, $y(0) = k_0$ is determined only within a discrete ambiguity. However, by examining the hits in the original data at $z = 0$ and at $z = 1$, the pair that differ by the fitted value of displacement k_1 can be identified as the beginning and end point of the track under question.

In the case of multiple tracks, the function $C(k)$ possesses D maxima in the two-dimensional space (k_1, k_2) .. Note that the maximum value of $C(k)$ is of order $\sim M$ if the tracks are nondegenerate. If D_d tracks are degenerate, i.e., have the same values of slope and curvature but *different* intercepts, then the maximum of $C(t)$ is of order $\sim D_d^2 M$. Thus these degeneracies can be estimated directly from the data and the values of $C(k)$ throughout the allowed region in k_1 and k_2 .

An initial estimate of the number of tracks and the values of k_1 can be made from a histogram of the function $C(k)$ against the scaled slope k_1 . Form the integral

$$C(k_1) = \int dk_2 \; C(k_1, k_2) \tag{45}$$

over the allowed range of values of k_2 . The peaks in k_1 that are of order M signify tracks. Two tracks that have the same value of k_1 but distinctly different values of k_2 produce a peak roughly twice as high. Two degenerate tracks with the same value of both k_1 and k_2 will produce a peak roughly four times as high. This initial survey of the data will simplify the search for all the relevant maxima of $C(k_1, k_2)$.

Now choose k_1 equal to one of the peak values of the histogram, say K_1 . Perform a one-dimensional search in k_2 of $C(K_1, k_2)$ for a peak, located at K_2 . This simple low dimensional search procedure could be continued by alternating directions to locate the position of the maxima. However it is convenient at this point to invoke the Simplex method. The required three starting simplexes are then initialized to the neighborhood of this approximate maxima. Using these as starting value, the "Uphill" Simplex method then searches the two dimensional space until the nearby (if our search was accurate) maxima is located. This maximum point then yields an estimate for all three k_i parameters describing one track. This process is repeated until there are no more maxima of $C(k)$ which are of magnitude M, i.e., large enough to be true tracks.

Alternatively, as each track is located and fitted, it can be subtracted from the data, i.e. the quantities c and s, before the next track is analyzed. As the detection efficiency drops, this method will eventually fail. A further (calculationally intensive) possibility is to return to the original pixel data, eliminate the 'hits' from the data that belong to the fitted track, and then repeat the entire process with the reduced data set.

In the case of multiple tracks, the determination of the values of k_1 and k_2 for each track is improved by using large values of $\delta\Phi$, i.e. short wave lengths, to resolve the differences between the tracks. This, of course, worsens the branch ambiguity in the value of the intercept k_0. Recalling eqns(22) and (38), and resolving the ambiguity in the arctangent function by requiring a maximum of the likelihood function, the final result for k_0 takes the form

$$\delta\Phi\, k_d(0) \;=\; -\arctan[\Sigma_s(d)/\Sigma_c(d)] + \; 2\, n\pi \; . \qquad (46)$$

The final ambiguity in $k_d(0)$ is in steps of $2\pi/\delta\Phi$. Thus the fitting procedure yields a discrete series of possible values for the intercept k_0. The correct value can be inferred by rerunning the program at an incommensurate value of $\delta\Phi$ and finding the common allowed value of $k_d(0)$. Alternatively, one may examine the original data set, armed with the fitted values of k_1 and k_2 for every track, looking for the first and last pair of hits with the displacement value k_1.

An example application of the procedure is illustrated in Figure 3 which plots four tracks whose parameters are given in the first columns of Table 1. The interaction region was set just off the lower left corner of the plot. Note that this sneaky choice reduces the problem of determining the correct branch of the intercept k_0 .

Table 1 – Four Tracks – Ideal

input			fitted		
k0	k1	k2	k0	k1	k2
0.045	0.100	1.40	0.0108	0.101	1.43
0.012	0.800	0.40	0.0121	0.799	0.40
0.009	0.300	0.80	0.0115	0.299	0.78
0.001	0.600	1.20	0.0022	0.598	1.20

In this example, $\delta\Phi$ was equal to 200 and $M = 101$. There were no noise hits added and the detection efficiency was 100%. The histogram function $C(k_1)$ is shown in Figure 4. The top drawing plots the histogram from the original data set. After the fit to the first track has been subtracted, the histogram is recalculated on the modified data. This is plotted on the left middle. The process is continued until the values drop below the assigned threshold value. The branch uncertainty in the determination of k_0 is 0.0314 . Note that the first track has a k_0 value that is off by one cycle, that is, $k_0 = 0.0108 + 0.0314 \sim 0.0422$, which is reasonably close to the input value of 0.045 .

In the next example, the noise and a finite detection efficiency were included. Noise was added by randomly choosing one fifth of the m values and adding a noise hit uniformly distributed in y between zero and one. Inefficiency was included by randomly omitting one fifth of the data points. The results fluctuate somewhat from run to run; typical values are given in Table 2.

Table 2 – Four Tracks – Noise					
input			fitted		
k0	k1	k2	k0	k1	k2
0.045	0.100	1.40	0.011	0.102	1.41
0.012	0.800	0.40	0.014	0.798	0.38
0.009	0.300	0.80	0.010	0.301	0.78
0.001	0.600	1.20	0.002	0.601	1.20

The histogram functions $C(k_1)$ for this case are shown in Figure 5. Note that the peak values have dropped and the background has increased relative to those in Figure 4; however the maxima are still distinct.

In the final example, the effects of track parameter degeneracy was explored. In Figure 6 the four tracks whose parameters are given in the first columns of Table 3 and Table 4 are plotted. Table 3 lists the values for no noise and perfect efficiency while Table 4 includes the effects of the noise and efficiency levels used in Table 2.

Table 3 – Four Tracks – Degeneracy					
input			fitted		
k0	k1	k2	k0	k1	k2
0.01	0.100	1.20	0.005	0.100	1.29
0.01	0.100	1.00	0.014	0.100	0.90
0.01	0.100	0.80	0.014	0.100	0.74
0.01	0.100	0.10	0.012	0.100	0.08

Table 4 – Four Tracks – Degen+Noise					
input			fitted		
k0	k1	k2	k0	k1	k2
0.01	0.100	1.20	0.005	0.100	1.28
0.01	0.100	1.00	0.015	0.098	0.90
0.01	0.100	0.80	0.014	0.101	0.75
0.01	0.100	0.10	0.013	0.100	0.06

Note that the k_1 slope parameters were accurately fitted, the k_2 parameters were determined with less accuracy, and the k_0 intercepts have large fractional errors.

5 Conclusions

The track fitting method developed here seems to offer some advantages in actual implementation. For the analysis of many events in the same detector, which is the normal situation in high energy physics experiments, many of the quantities can be precomputed and stored for use during an event by event analysis. Efficient algorithms exist for locating (with the required accuracy) the maximum of $C(t)$ in the low dimensional track parameter space. Clearly, further testing of this algorithm in more realistic situations is required.

Acknowledgments

I wish to thank Professor T. Kailath for a helpful discussion and Professor Sid Drell for useful suggestions.

A Antenna Arrays

It is the purpose of this appendix to map the problem of fitting multiple tracks, or lines, to the problem of determining the directions of arrival of waves incident upon an antenna array. First, the antenna problem will be stated. The discussion will be restricted to two dimensions for simplicity; three dimensional tracks can always be projected onto lower dimensions. Only straight line tracks will be discussed here. Extensions to tracks with curvature is given in the text.

Antenna-Arrival Directions

Consider a straight line sensor array consisting of M antenna elements aligned along the z-axis. The location of the m^{th} sensor is denoted by z_m . A pure harmonic plane wave of constant amplitude is incident upon the array, where

$$s(t) \ = s \exp[-i\omega t] \ , \qquad s \ = \rho \exp[i\phi] \ . \tag{47}$$

The sensors are characterized by the array response vector, which contains the phase lag at each sensor, given by

$$\mathbf{a} \;=\; [\ a_0(\theta),\ a_1(\theta), ...,\ a_{M-1}(\theta)\ ]\ , \tag{48}$$

where $a_m(\theta)$ $(0\ <= m\ <\ M-1)$ is the amplitude induced at the m^{th} sensor by a *unit* plane wave arriving from the direction θ . The collection of all the response vectors over the range of interest in theta is termed the *array manifold*.

Choosing the arbitrary phase of the wave at the 0^{th} sensor to vanish, the elements of the array response vector are given by

$$a_m(\theta) \;=\; \exp[iz_m \sin\theta]\ , \tag{49}$$

with $a_0(\theta) \equiv 1$ and $z_0 \equiv 0$. The m^{th} element of the output vector $\mathbf{o}(t)$ is the response of m^{th} sensor to the incident wave; it is given by

$$o_m(t) \;= a_m(\theta)\ s(t)\ . \tag{50}$$

Antenna-Multiple Sources

Now consider the case of D sources whose waves arrive at the array from different angles. The wave from the d^{th} , $(0\ <= d\ <\ D-1)$, source is

$$s_d(t) \;= s_d\ \exp[-i\omega t]\ , \qquad s_d \;= \rho_d \exp[i\phi_d]\ . \tag{51}$$

Thus the output at the m^{th} sensor is the sum over the sources

$$o_m(t) \;=\; \sum_0^{D-1} a_m(\theta_d)\ s_d(t)\ , \tag{52}$$

This can be written as a matrix equation by forming a D-component column vector out of the s_d's together with a matrix $\mathbf{A}$ of M columns and D rows. Each row is formed from the M-component vector $\mathbf{a}(\theta_d)$. Then one can write

$$\mathbf{o}(t) \;=\; \mathbf{A}(\theta)\mathbf{s}(t)\ . \tag{53}$$

Armed with this review, the discussion will now switch to track fitting.

B Linear Track Fitting

Assume that the data for the d^{th} track reflects hits that are along a line

$$x_m(d) \;=\; x_0(d) \;+\; z_m \;\tan\theta_d \;, \tag{54}$$

where $x_0(d)$ is the intercept and $\tan\theta_d$ is the slope of the track. We have also assumed perfect detection efficiency and no noise hits. Now formally define 'signals' given by the sum over all the D tracks as

$$\begin{aligned} o_m \;&=\; \sum_{d=0}^{D-1} \exp[ikx_m(d)] \;=\; \sum_{d=0}^{D-1} \exp[ikz_m \;\tan\theta_d] \cdot \exp[ikx_0(d)] && (55)\\ &\equiv\; \sum_{d=0}^{D-1} a_m(\theta_d) \cdot s_d \;, && (56)\end{aligned}$$

where k is a parameter to be chosen later for convenience and we have introduced the quantities

$$s_d \;=\; \exp[ikx_0(d)] \qquad \text{and} \qquad a_m(\theta) \;=\; \exp[ikz_m \;\tan\theta] \;. \tag{57}$$

Now form the discrete Fourier transform $O(t)$ of the signal vector $\mathbf{o}$ with the transform variable scaled by k :

$$\begin{aligned} O(t) \;&=\; \sum_{m=0}^{M-1} o_m \;\exp[-iktz_m] = \sum_{d=0}^{D-1} \; s_d \sum_{m=0}^{M-1} a_m(\theta_d) \;\exp[-iktz_m] && (58)\\ &=\; \sum_{d=0}^{D-1} s_d \; \sum_{m=0}^{M-1} \exp[iz_m k(\tan\theta_d - t)] \;. && (59)\end{aligned}$$

As a function of the scaled transform variable t , the function $O(t)$ has a maximum when $t \sim \tan\theta_d$.

This is easily illustrated if the array has uniform spacing, $z_m = m\delta z$. The sum over m can then be performed in closed form with the result

$$O(q) \;=\; \sum_{d=0}^{D-1} s_d \;\exp[i(M-1)\Delta_d] \;\frac{\sin(M\Delta_d)}{\sin(\Delta_d)} \;, \tag{60}$$

where $\Delta_d = \frac{1}{2}k\delta z(\tan\theta_d - t)$. The function $O(t)$ has a maximum whenever Δ_d vanishes. This will eventually allow the determination of the angles θ_d for all d . For example, if the tracks are well separated in angle, then as the parameter t is varied the real part will have a maximum at $t = t_e = \tan\theta_e$. For this value of t the output signal is

$$\begin{aligned} O(t_e) &= Ms_e + \sum_{d\neq e} s_d \exp[i(M-1)\Delta_d] \frac{\sin(M\Delta_d)}{\sin(\Delta_d)} && (61) \\ &\sim Ms_e + O(1) . && (62) \end{aligned}$$

The other tracks will not yield contributions of order M due to oscillations; the value of $k\delta z$ is chosen to insure this cancellation. This result also allows a lowest order estimate of the intercept from $s_e \sim O(t_e)/M$.

Note that the track fitting problem has been transformed into an antenna problem with the simple replacement of $\tan\theta$ by $\sin\theta$. The reason for this is that in the antenna problem, the waves travel in a direction perpendicular to the wave front. The track fitting problem has 'waves' that move perpendicular to the z-direction; the resultant phase lags are therefore different functions of the angle.

Degeneracy

If two tracks, say e and f , have essentially the same angle, then the sum becomes

$$\begin{aligned} O(t_e) &= M(s_e + s_f) + \sum_{d\neq e,f} s_d \exp[i(M-1)\Delta_d] \frac{\sin(M\Delta_d)}{\sin(\Delta_d)} && (63) \\ &\sim M(s_e + s_f) + O(1) . && (64) \end{aligned}$$

In this case, the absolute square of $O(t)$ becomes

$$|O(t_e)|^2 \sim M^2|s_e + s_f|^2 = 2\,M^2\{1 + \cos(k[x(e) - x(f)])\}, \qquad (65)$$

and the magnitude depends upon the relative phase, that is, the distance between the parallel tracks.. By studying the variation with k , the presence of degenerate angles can be inferred from the magnitude compared to M^2 , and the displacement between the tracks can be estimated.

References

[1] H. K. Aghajan and T. Kailath, "SLIDE: Subspace-Based Line Detection," *IEEE Trans. Pattern Anal. Machine Intell.* **16**, 1057-1073 (1994) and "Sensor Array Processing Techniques for Super Resolution Multi-Line-Fitting and Straight Edge Detection," *IEEE Trans. Signal Processing* **2**, 454-465 (1993). An extensive list of earlier references are given in these papers.

[2] R. Roy and T. Kailath, "ESPRIT–Estimation of Signal Parameters Via Rotational Invariance Techniques,"*IEEE Trans. Acoust., Speech, and Signal Processing* **37**, 984-995 (1989). See also R. Roy, A. Paulraj and T. Kailath, "ESPRIT–A Subspace Rotation Approach to Estimation of Parameters of Cisoids in Noise,"*IEEE Trans. Acoust., Speech, and Signal Processing* **ASSP-34**, 1340-1342 (1986).

[3] X-M. Lou, L. G. Hassebrook, M. E. Lhamon, and J. Li, "Numerically Efficient Angle, Width, Offset, and Discontinuity Determination of Straight Lines by the Discrete Fourier-Bilinear Transformation Algorithm," *IEEE Trans. on Signal Processing* **6**, 1464-1467 (1997).

[4] R. Kumaresan and D. Tufts, "Estimating the Angles of Arrival of Multiple Plane Waves," *IEEE Trans. Aero. and Elect. Sys.* **AES-19**, 134-139 (1983).

[5] D. C. W. Pao, H. F. Li, and R. Jayakumar, "Shapes Recognition Using the Straight Line Hough Transform: Theory and Generalization," *IEEE Trans. Pattern Anal. Machine Intell.* **14**, 1076-1089 (1992). Earlier references can be found here.

[6] M. Ohlsson, C. Peterson, A. Yuille, "Track Finding with Deformable Templates—The Elastic Arms Approach," *Computer Physics Communications* **71**, 77 (1992). An extensive list of earlier references are given here.

[7] M. Ohlsson, "Extensions and Explorations of the Elastic Arms Algorithm," *Computer Physics Communications* **77**, 19 (1993).

[8] Richard Blankenbecler, "Deformable templates — revisited and extended — with an OOP implementation," *Computer Physics Communications* **81**, 318-334 (1994). See also "A unified treatment of track reconstruction and particle identification," *Computer Physics Communications* **81**, 335-342 (1994).

[9] A. Schuster, "On Lunar and Solar periodicities of Earthquakes," *Proc. Roy. Soc.* **61**, 455-465 (1897).

[10] E. T. Jaynes (1987), "Bayesian Spectrum and Chirp Analysis," in C. R. Smith and G. J. Erickson (ed.), *Maximum-Entropy and Bayesian Spectral Analysis and Estimation Problems* D. Reidel, Dordrecht.

[11] G. L. Bretthorst, *Bayesian Spectrum Analysis and Parameter Estimation* Springer-Verlag, New York (1988).

[12] G. J. Erickson, P. O. Neudorfer and C. R. Smith, p. 505-509 (1989), J. Skilling (ed.), *Maximum Entropy and Bayesian Methods* Kluwer Academic Publishers.

[13] If these approximations do not hold, then the value of $\delta\Phi$ is inappropriate, or M is not large to begin with. In the latter case, the following results lead to a first order answer which can be easily improved by returning to the exact sums.

[14] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling, *Numerical Recipes: the art of scientific computing.* Cambridge Univ. Press, 1986. 818p.

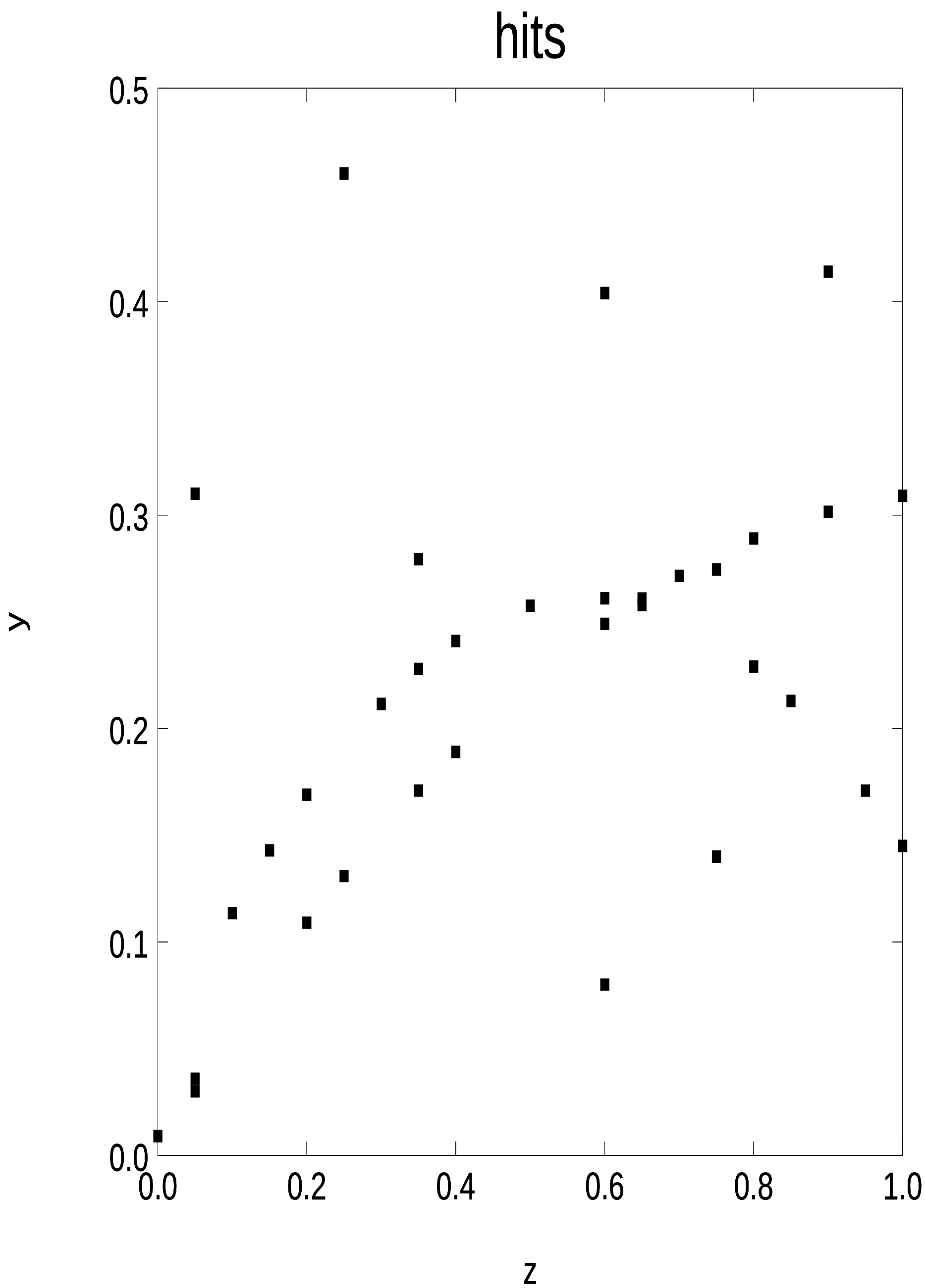

FIGURE 1
A Sample Event with Noise and Finite Detection Efficiency.

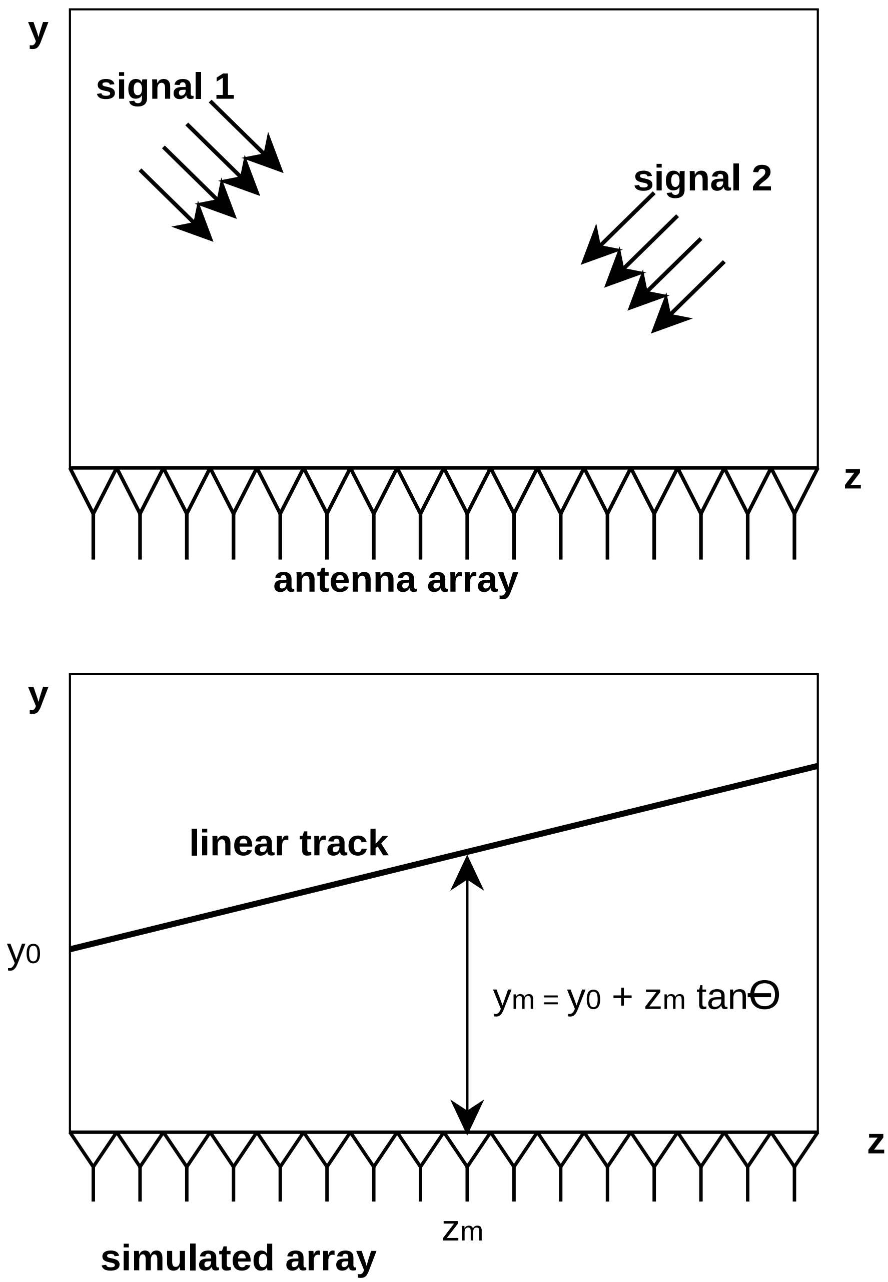

FIGURE 2
A Schematic of the Antenna problem and the Particle Track Analogue.

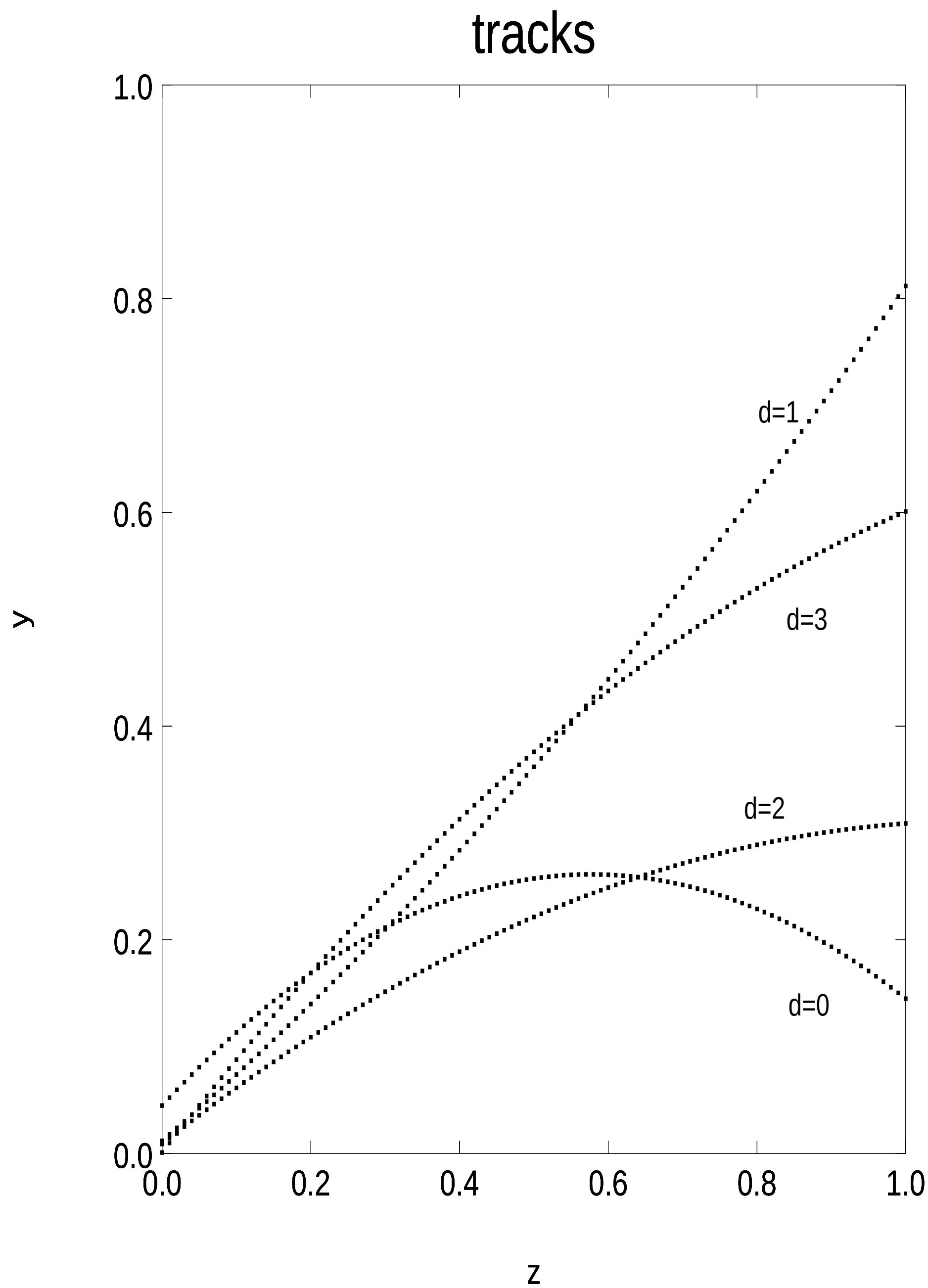

FIGURE 3
A Sample Event with 4 tracks.

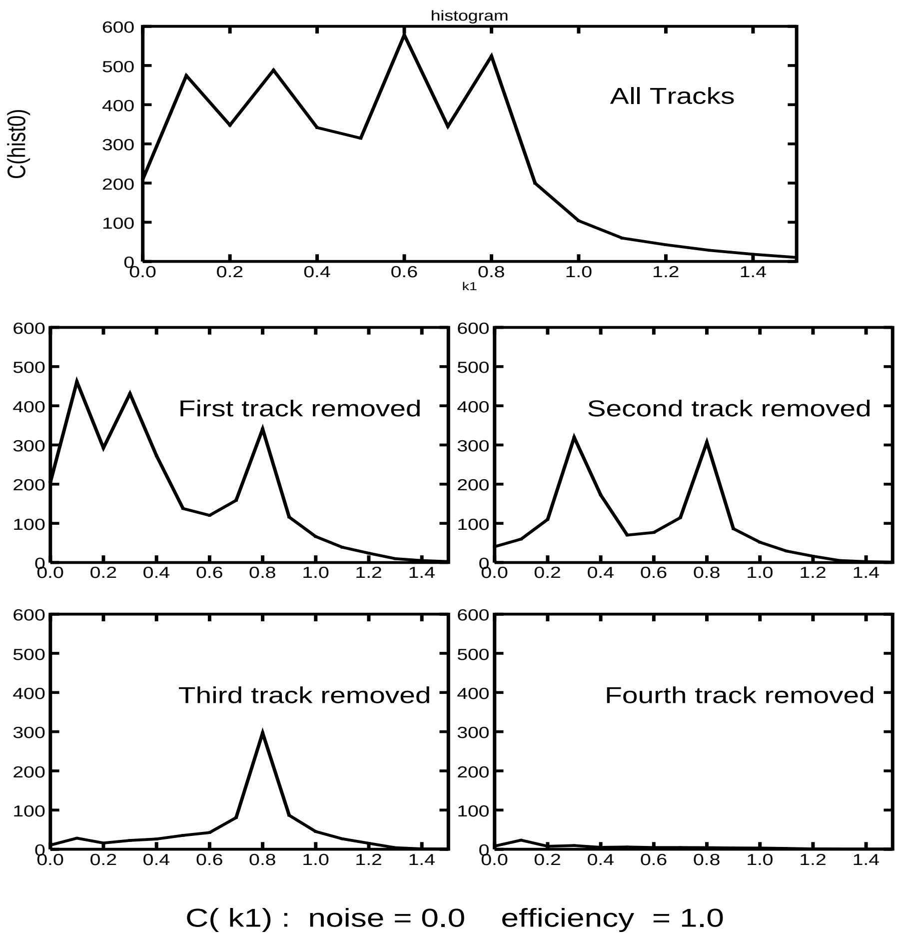

FIGURE 4

A plot of the histogram C(k_1) for zero noise and perfect detector efficiency as fitted tracks are removed.

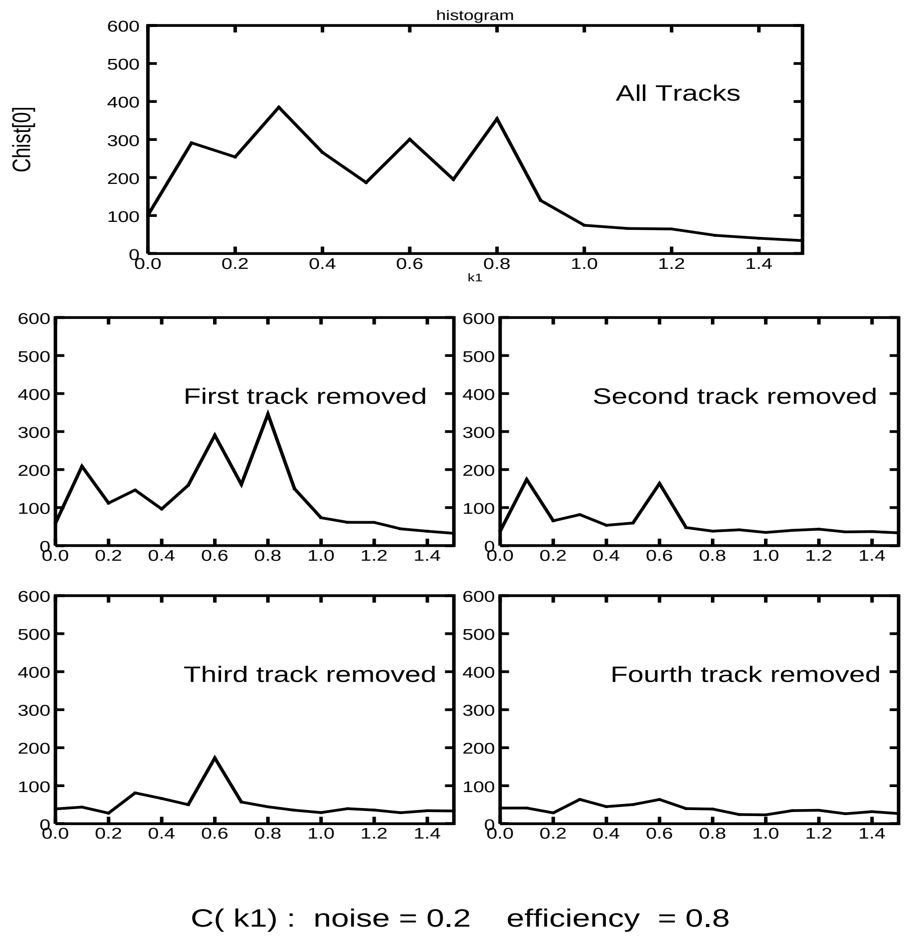

FIGURE 5

The same histogram plot as in previous figure but with finite noise and detector efficiency.

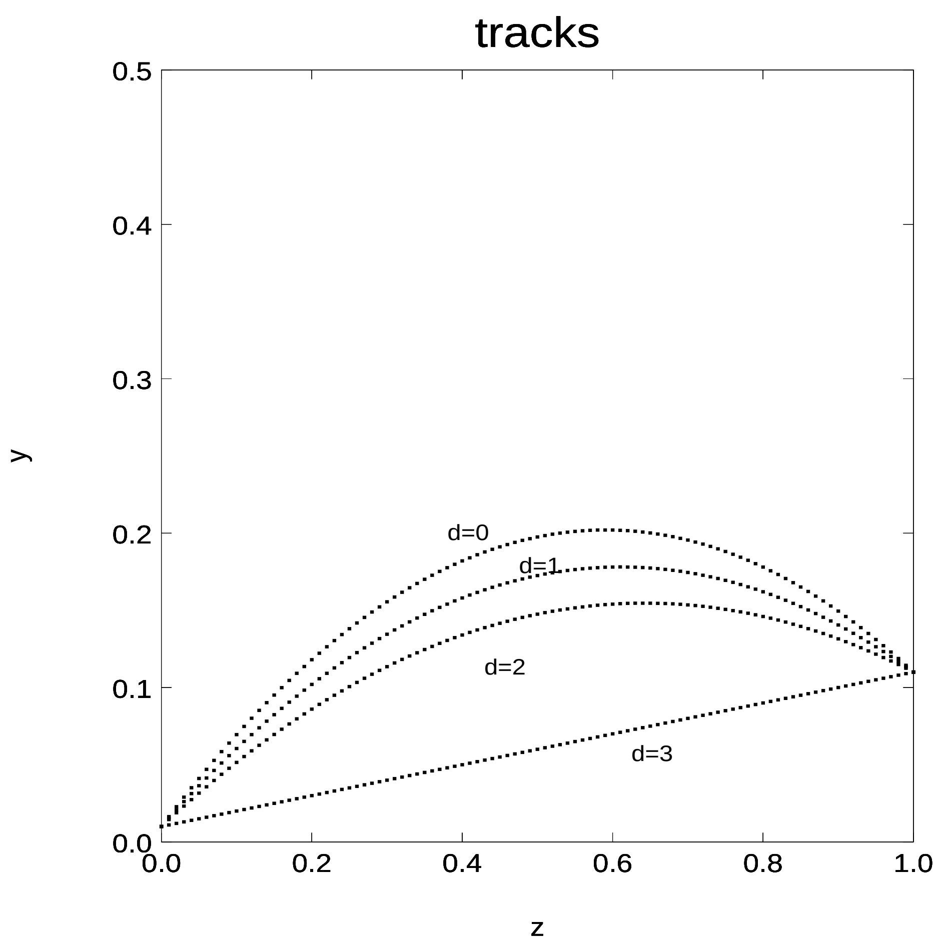

FIGURE 6
A Sample Event with 4 tracks with degeneracy.