

SCHOOL OF DATA ANALYSIS

Hough Transform + ML

Mikhail Hushchyn, Andrey Ustyuzhanin

Hough Transform + ML



Hough Transform for a Hit

In polar coordinates (r, ϕ) :

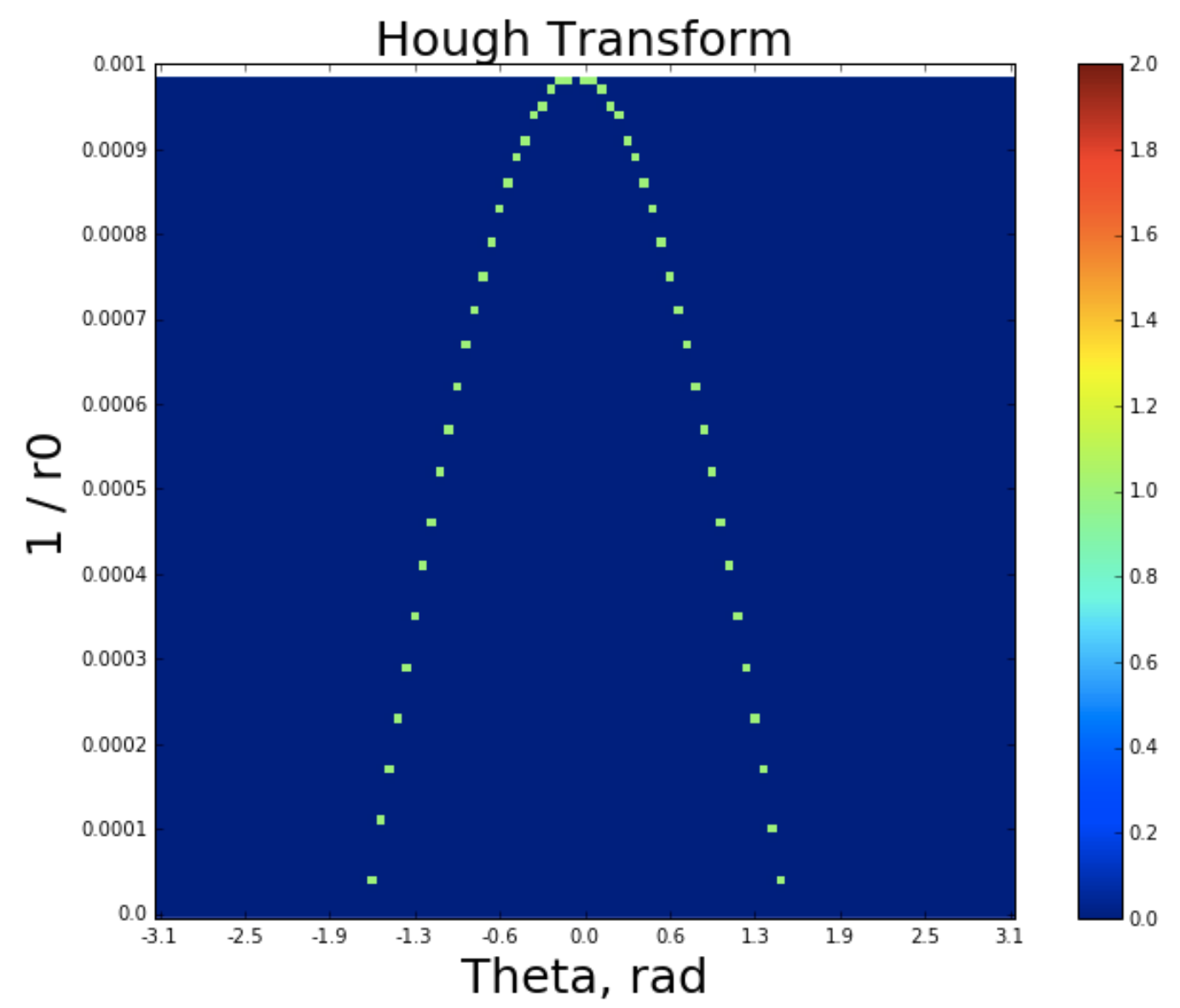
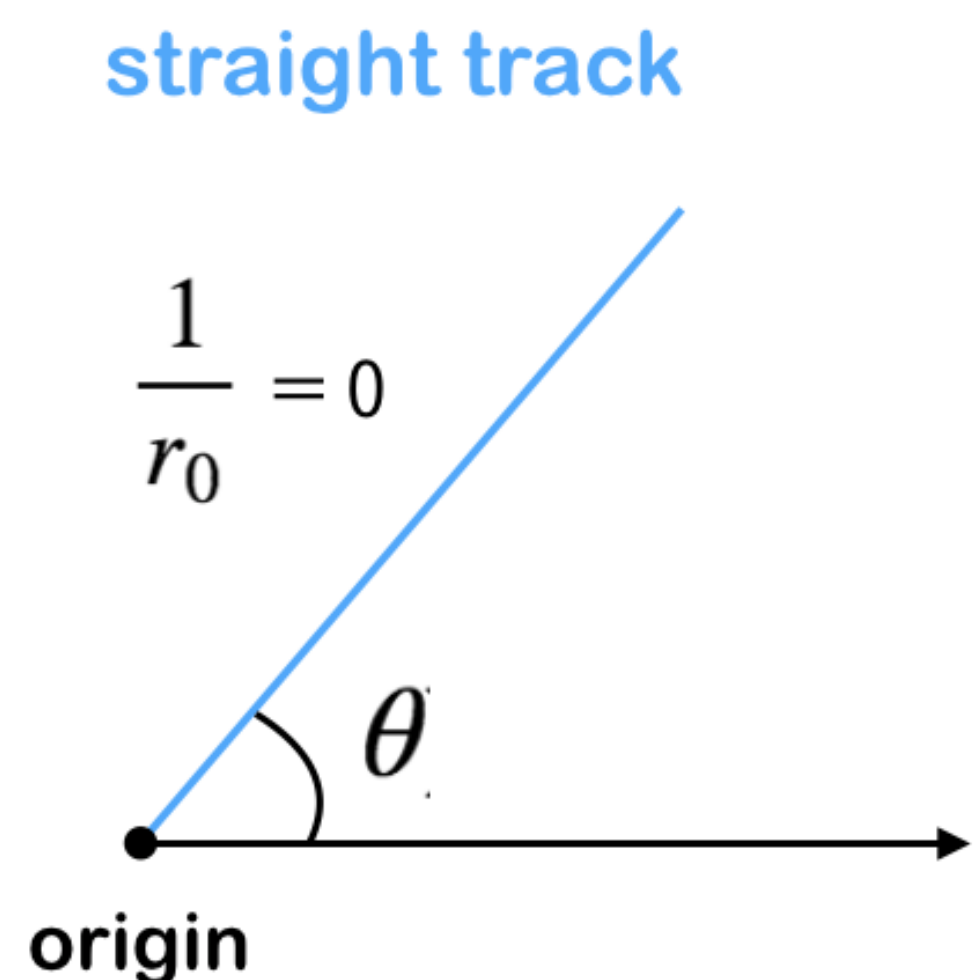
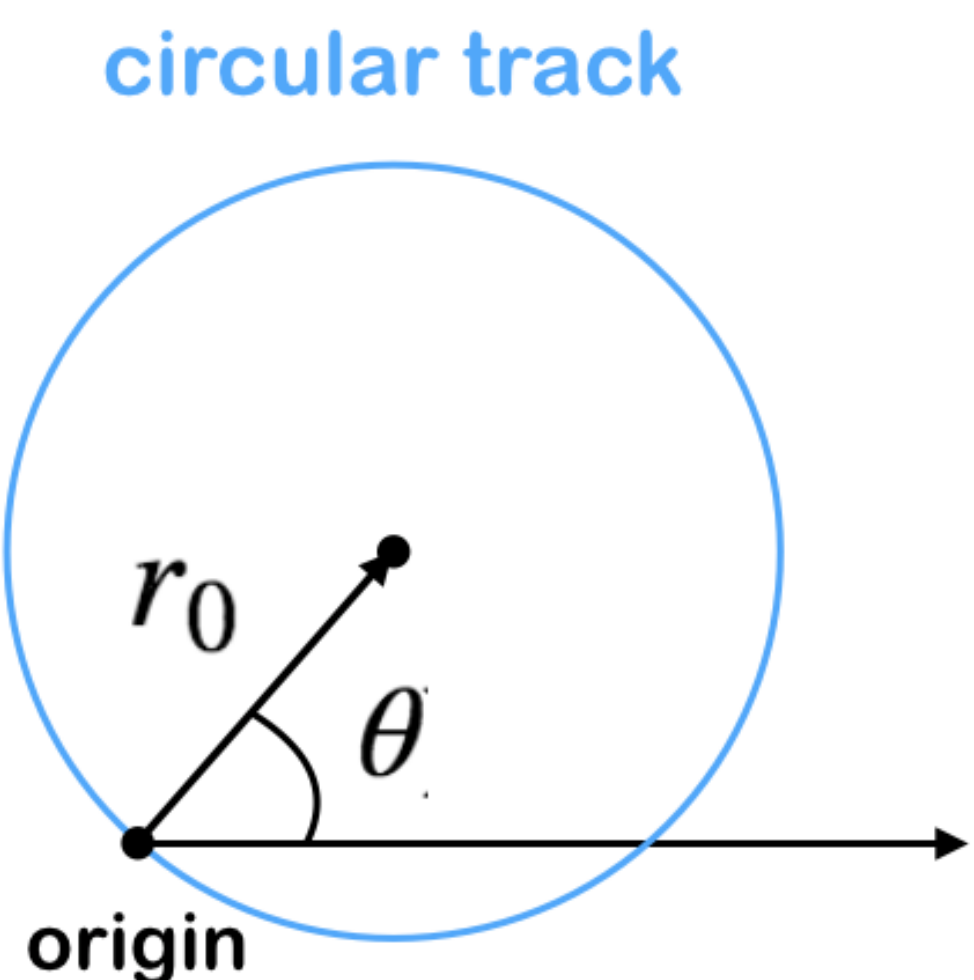
$$r = 2r_0 \cos(\phi - \theta)$$

One hit with coordinates (r, ϕ) :

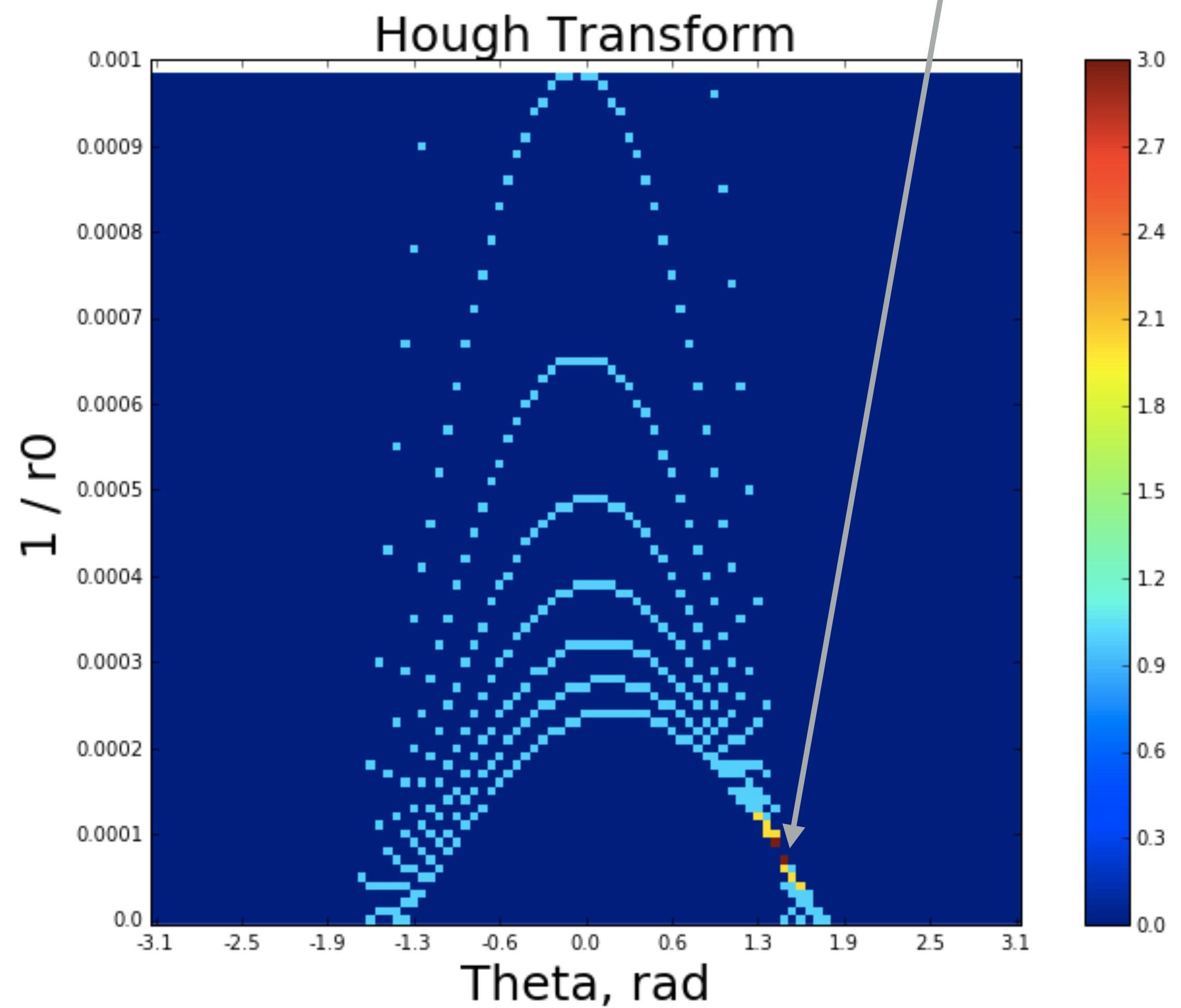
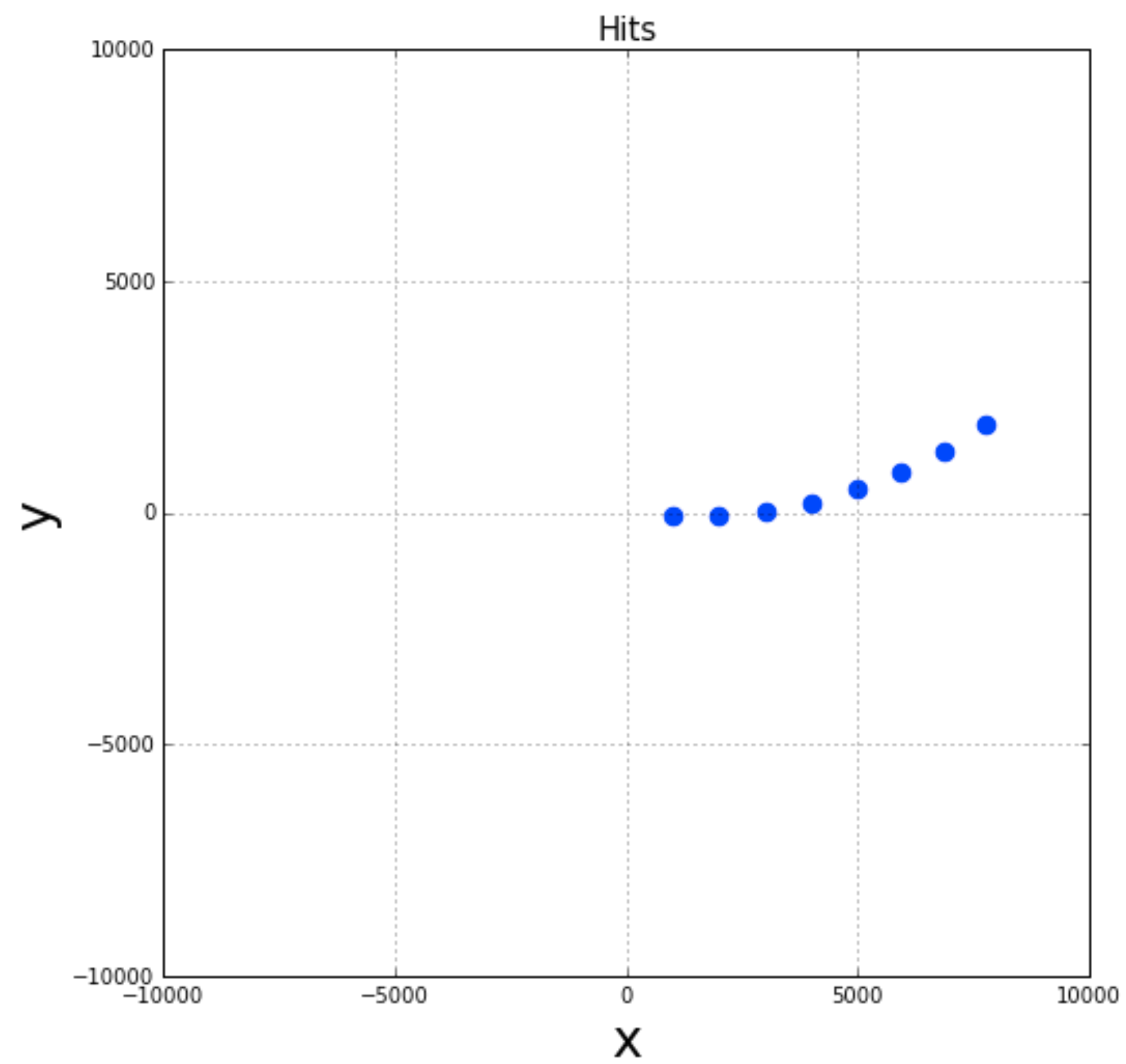
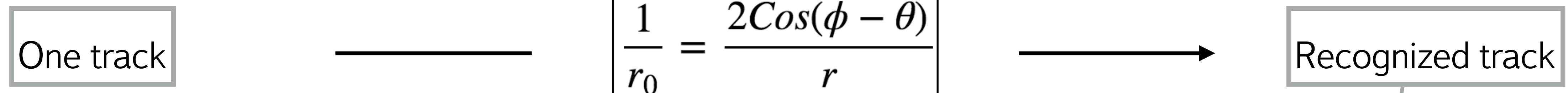
Hit (r, ϕ)

Transform

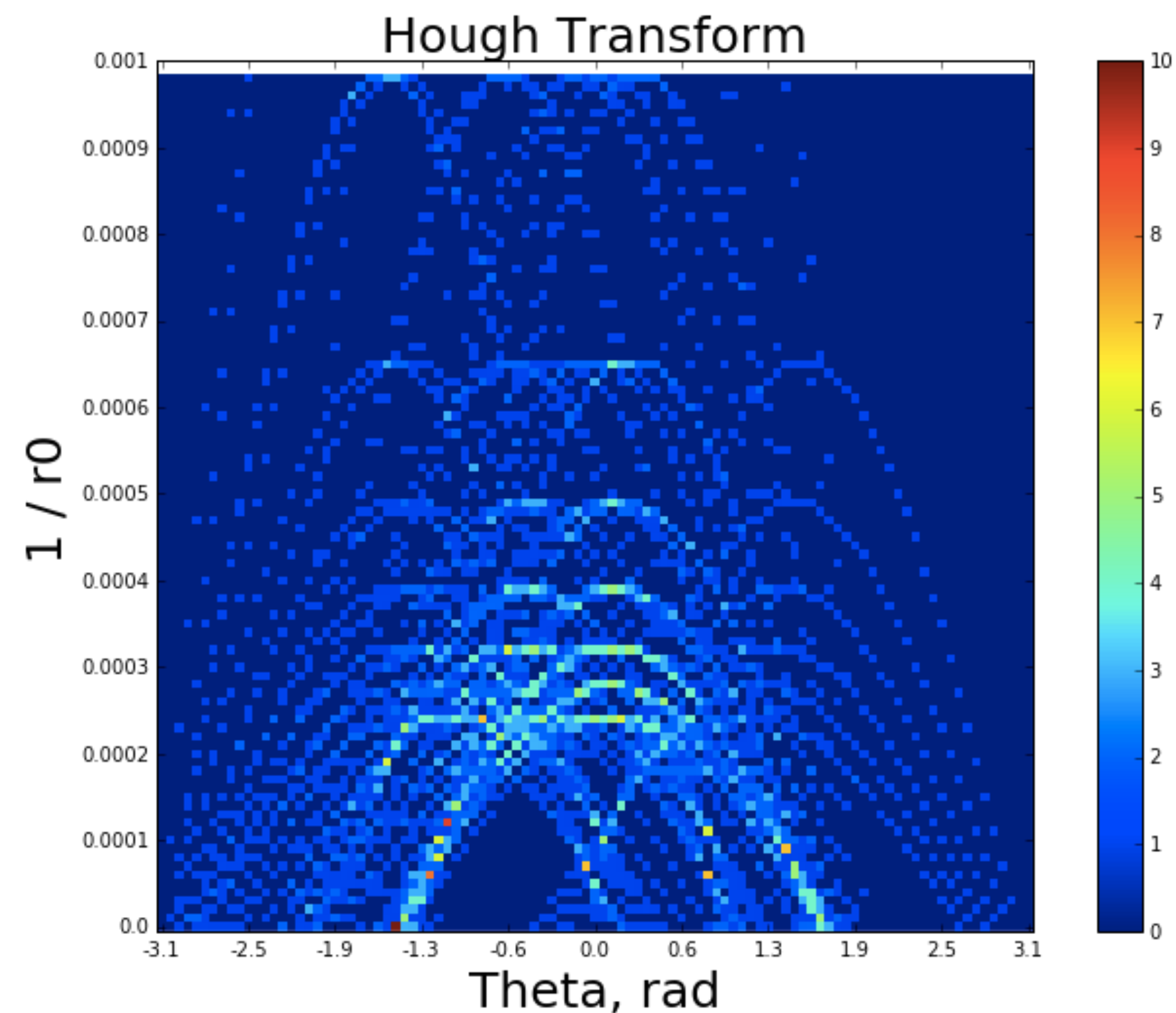
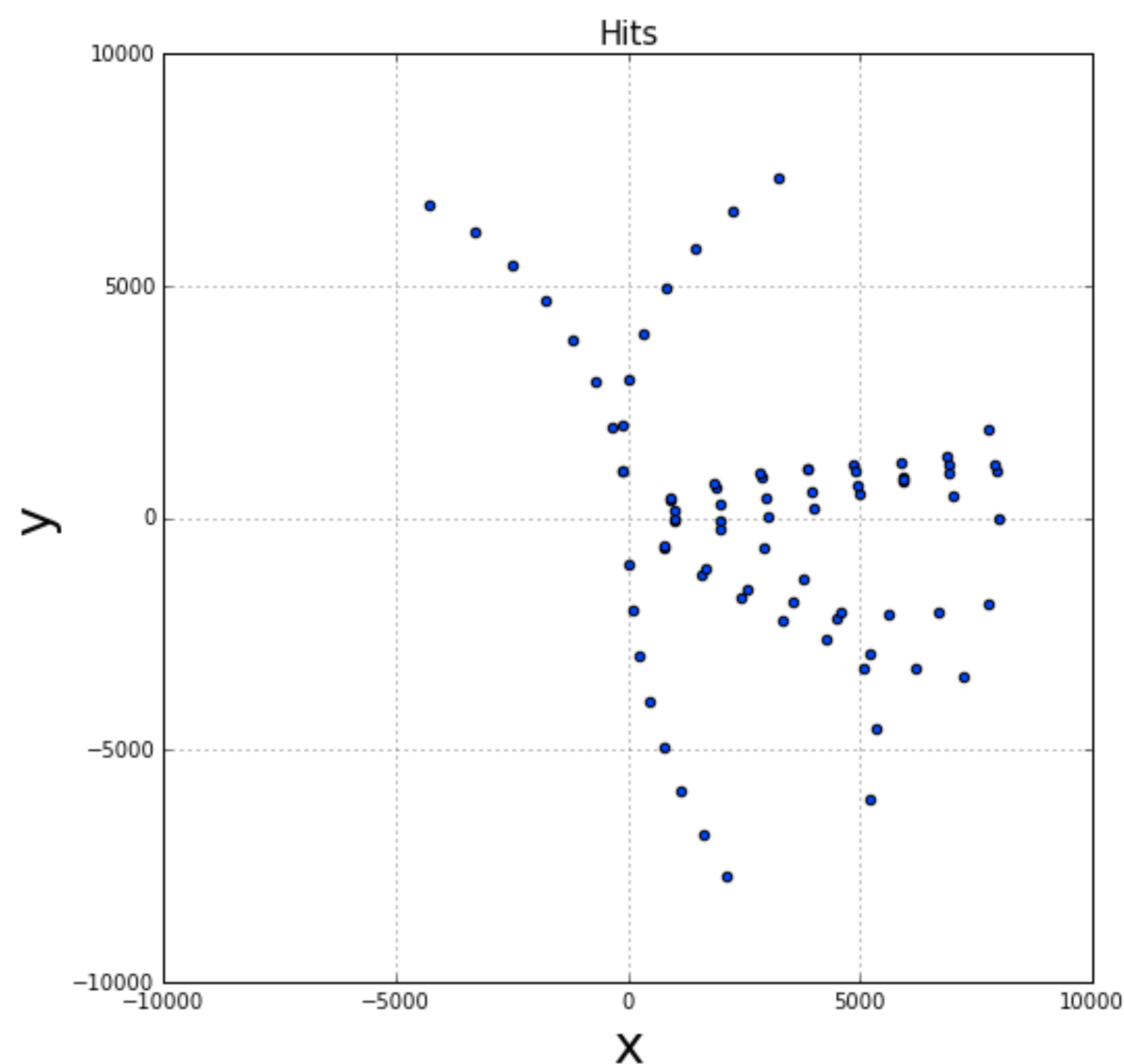
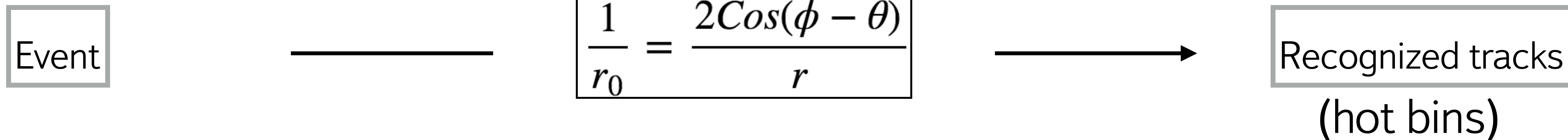
$$\frac{1}{r_0} = \frac{2 \cos(\phi - \theta)}{r}$$



Hough Transform for a Track



Hough Transform for an Event



Recognized tracks: good tracks, clones, ghosts.

Hough Transform + Tracks Clustering

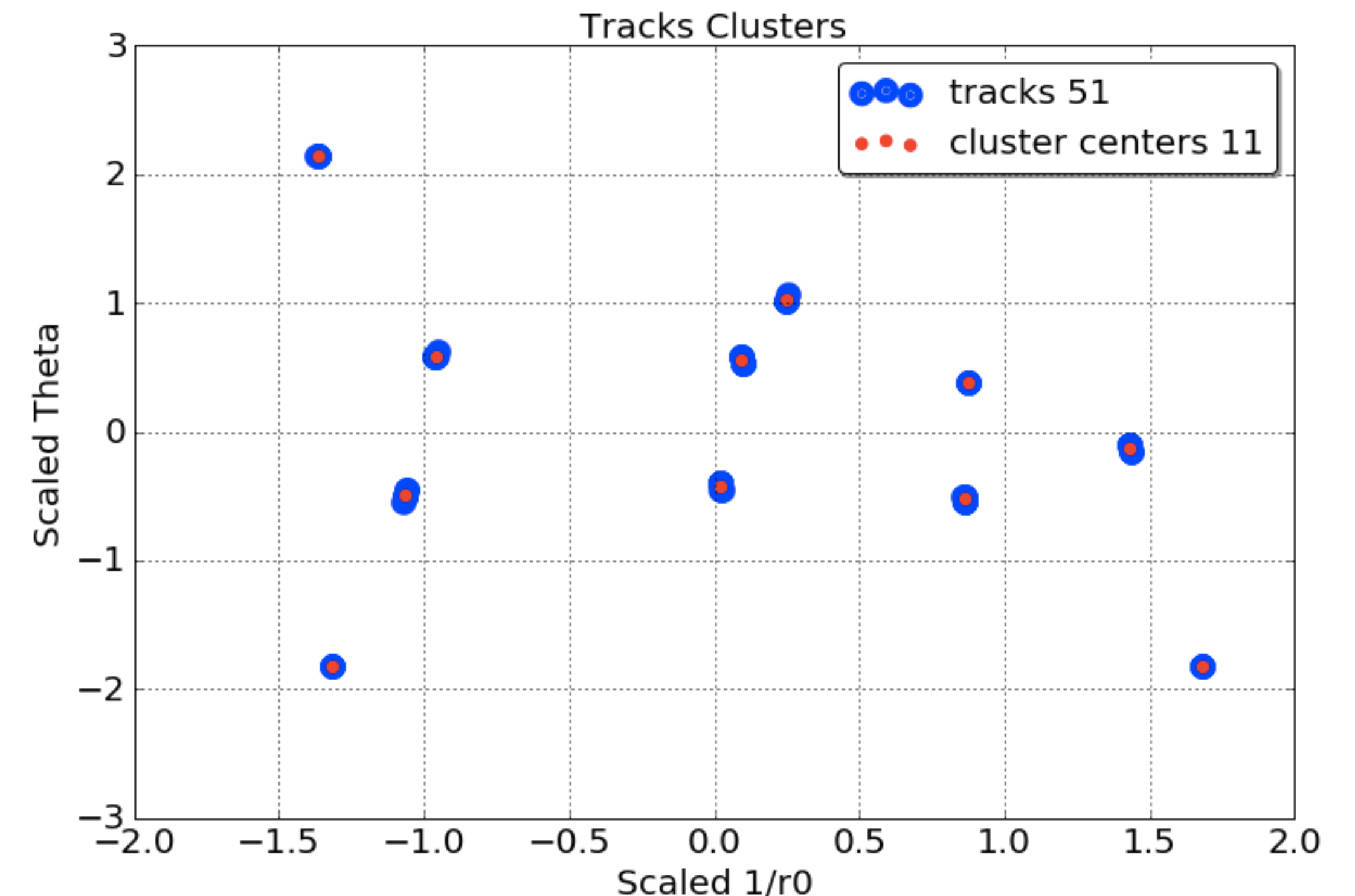
One event with 10 tracks:

Tracks clustering to reduce a number of clones.

Features: Track parameters

Methods: K-Means, Mean-shift, DBSCAN, Agglomerative clustering, ... ([more](#))

Metrics: Fowlkes-Mallows scores, Homogeneity, Completeness and V-measure, Silhouette Coefficient, ... ([more](#))



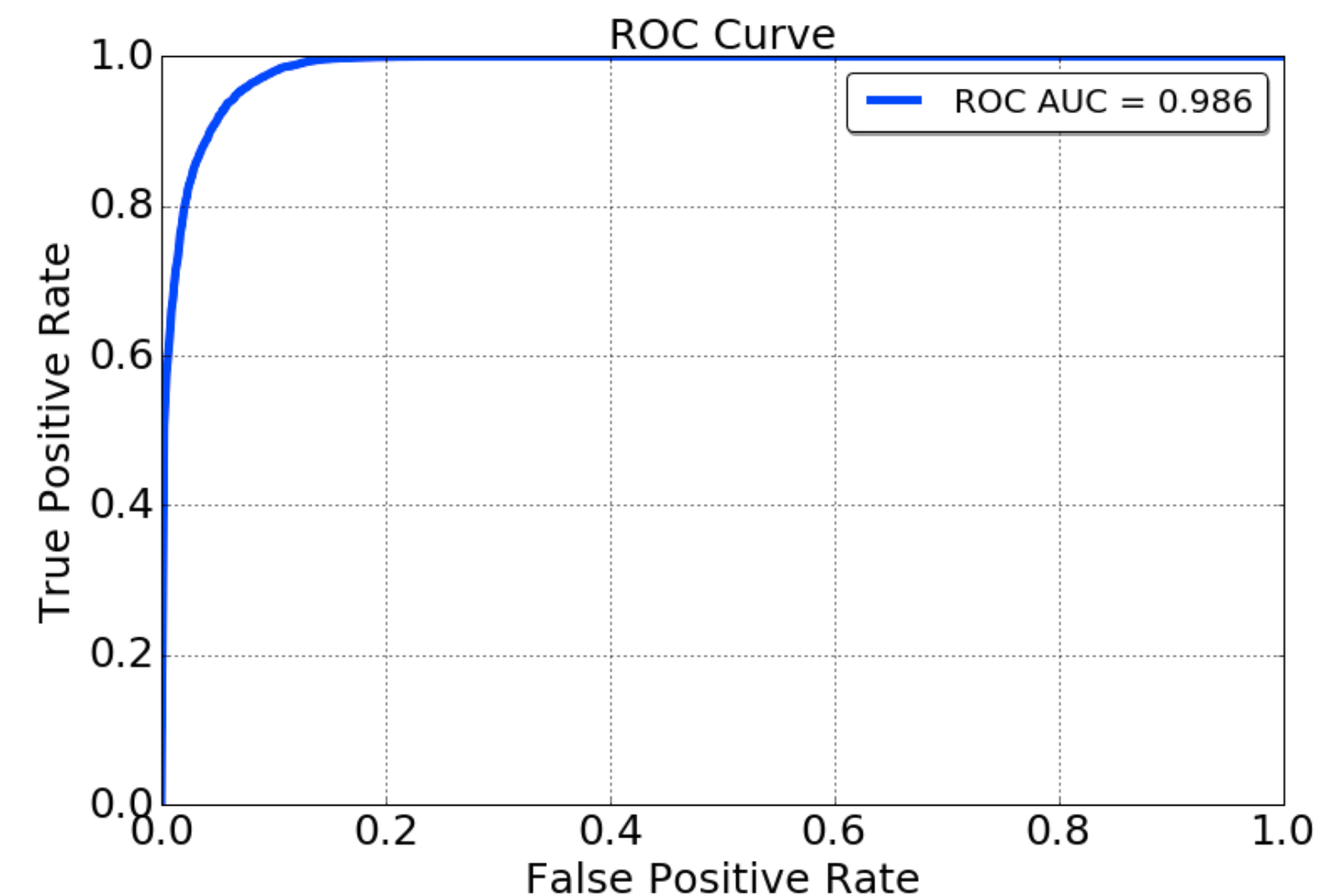
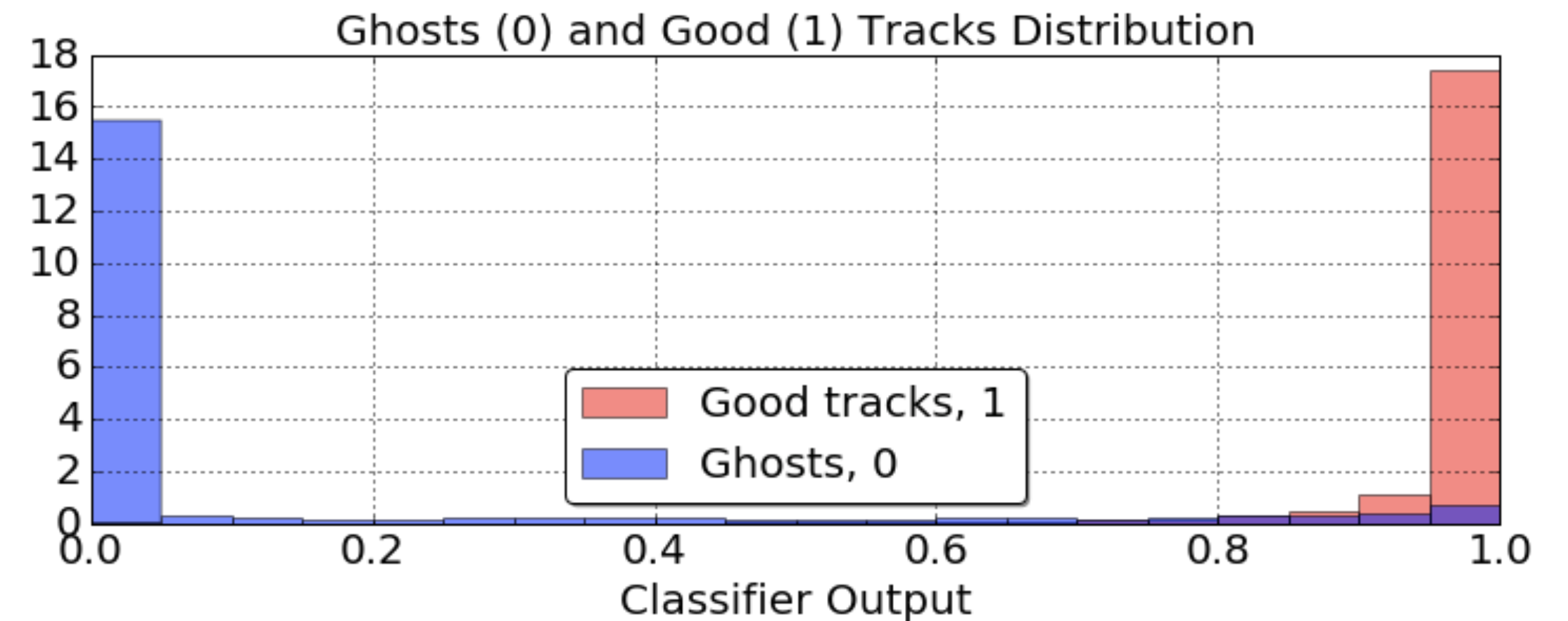
Hough Transform + Tracks Classification

Tracks classification to reduce a number of ghosts.

Features: Track parameters, number of hits, RMSE of a track fit

Methods: ANN, Random Forest, Gradient Boosting, ...

Metrics: ROC-curve, ROC AUC



Two approaches:

1) Each hot bin is a recognized track. This means, that one hit can belong to several recognized tracks:

reco. track 1: 1, 2, 3, 4, 5

reco. track 2: 4, 5, 6, 7, 8

...

2) One hit belongs to just one recognized track. This means, each hit has only one recognized track label:

Reco. hit labels: 1, 1, 1, 2, 2, 2, 3, ...

True hit labels: 1, 1, 1, 1, 2, 2, 2, ...

The 2nd approach goes from the 1st one. Not vice versa!