!!! TITLE

USING CLUSTERING TO IDENTIFY 3 MILLION NEW ASTEROIDS

PREPARED BY

MATT HOLMAN
PAUL BLANKLEY
RYAN JANSSEN

CS182 – ARTIFICIAL INTELLIGENCE

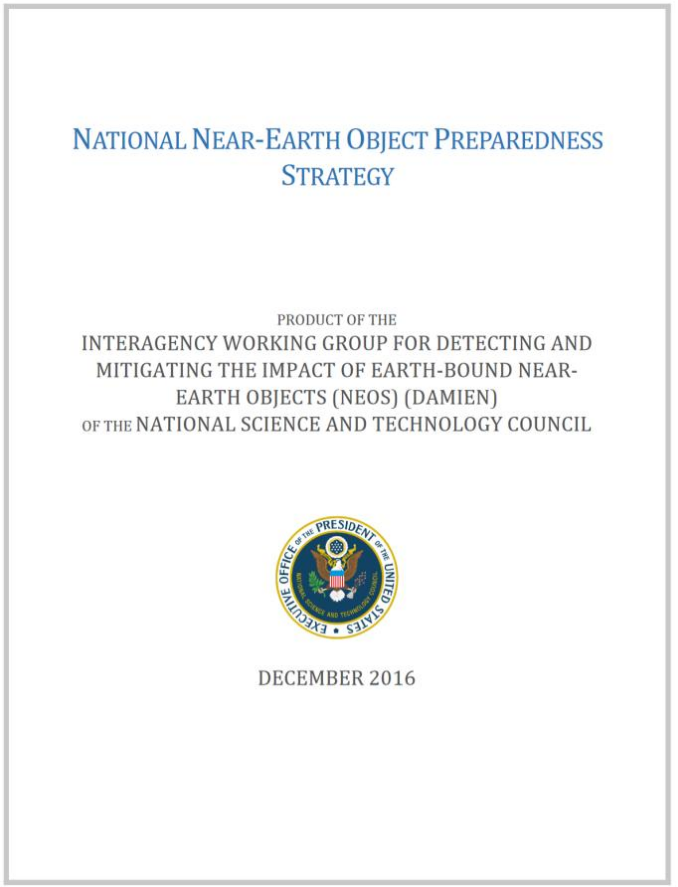
30 NOVEMBER 2017

MILLIONS OF MISSING ASTEROIDS!

CONNECTING PARTIAL ASTEROID TRACKLETS

Mandate from Congress:
>90% of potentially hazardous asteroids must be tracked

Currently, we have **170m observations** of known asteroids



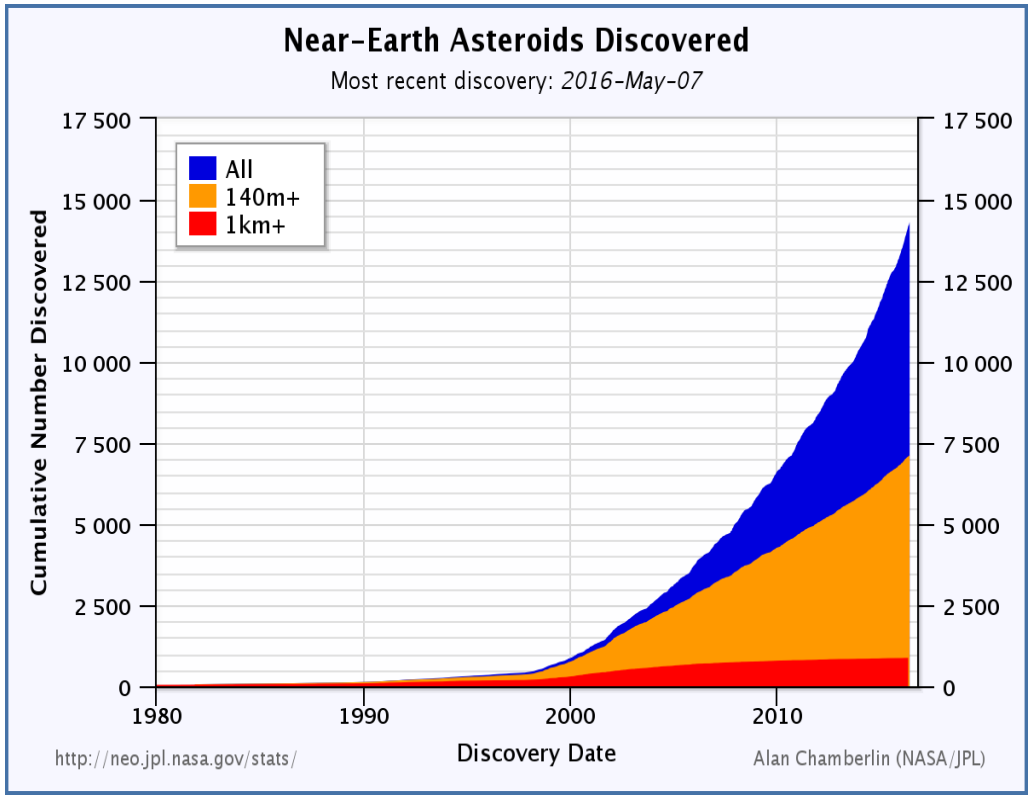
NATIONAL NEAR-EARTH OBJECT PREPAREDNESS STRATEGY

PRODUCT OF THE
INTERAGENCY WORKING GROUP FOR DETECTING AND MITIGATING THE IMPACT OF EARTH-BOUND NEAR-EARTH OBJECTS (NEOS) (DAMIEN)
OF THE NATIONAL SCIENCE AND TECHNOLOGY COUNCIL

DECEMBER 2016

MILLIONS OF MISSING ASTEROIDS!

CONNECTING PARTIAL ASTEROID TRACKLETS



However, some observations cannot be matched into an asteroid.

There are currently:

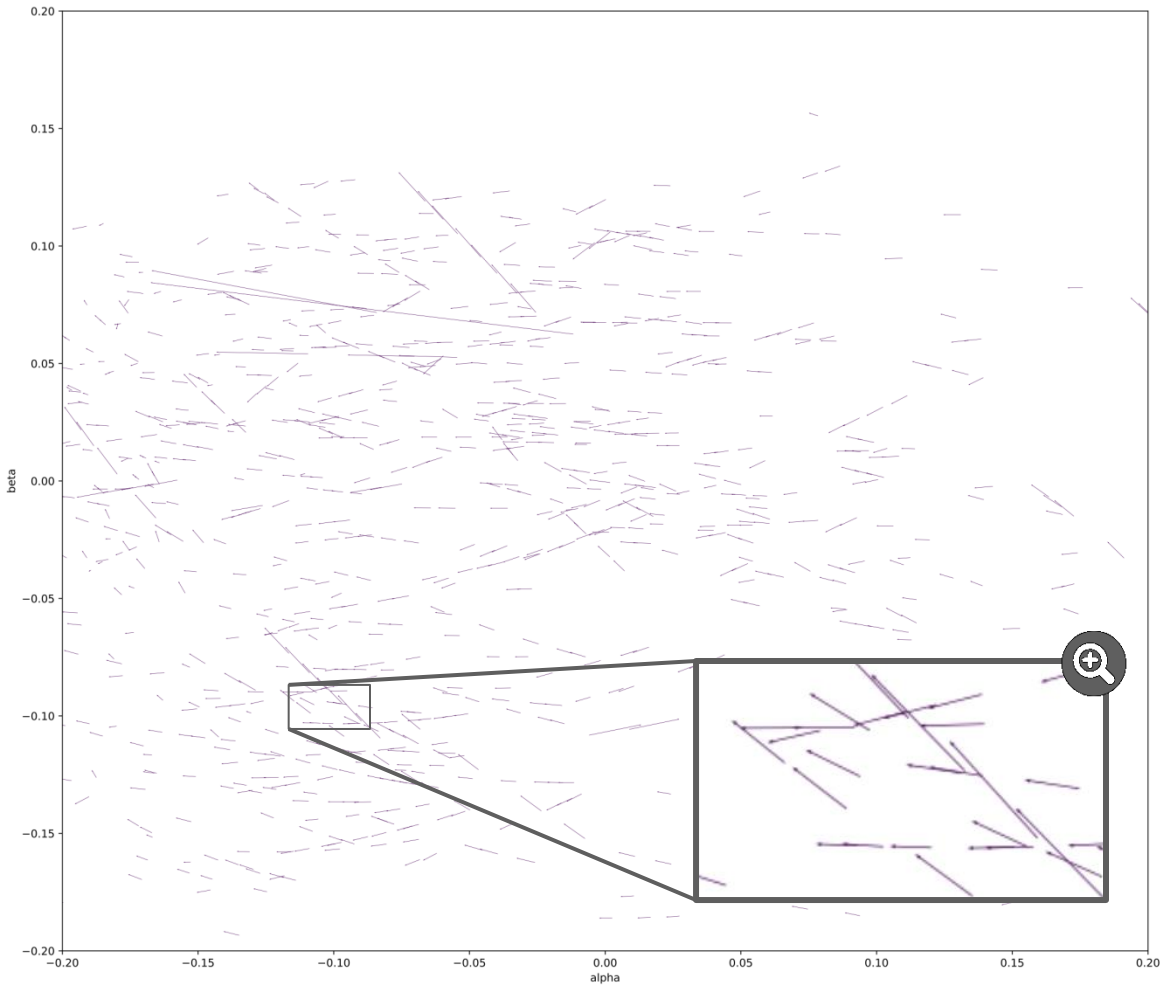
- **14 million** unmatched tracklets,
- **~3 million** total asteroids.

Growing quickly as telescope resolution increases

A HARD PROBLEM...

- High asteroid density makes clustering difficult
- From Earth, asteroids trajectories are highly non-linear
- Missing data (cannot tell asteroid movement/velocity in the z direction)
- Sparse observations can be months or years apart
- Brute force is impossible with 14m observations

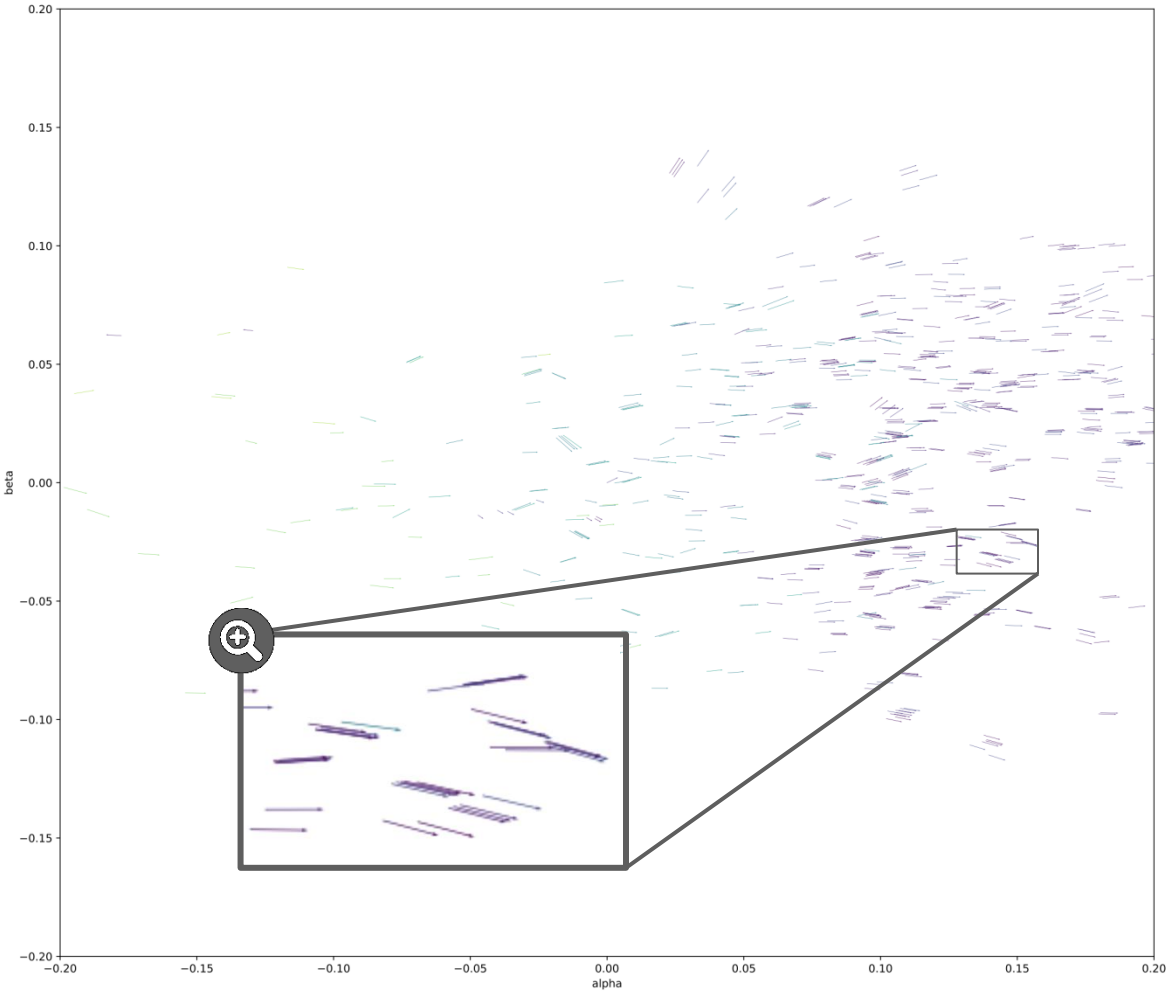
Partial asteroid tracklets (from Earth's perspective)



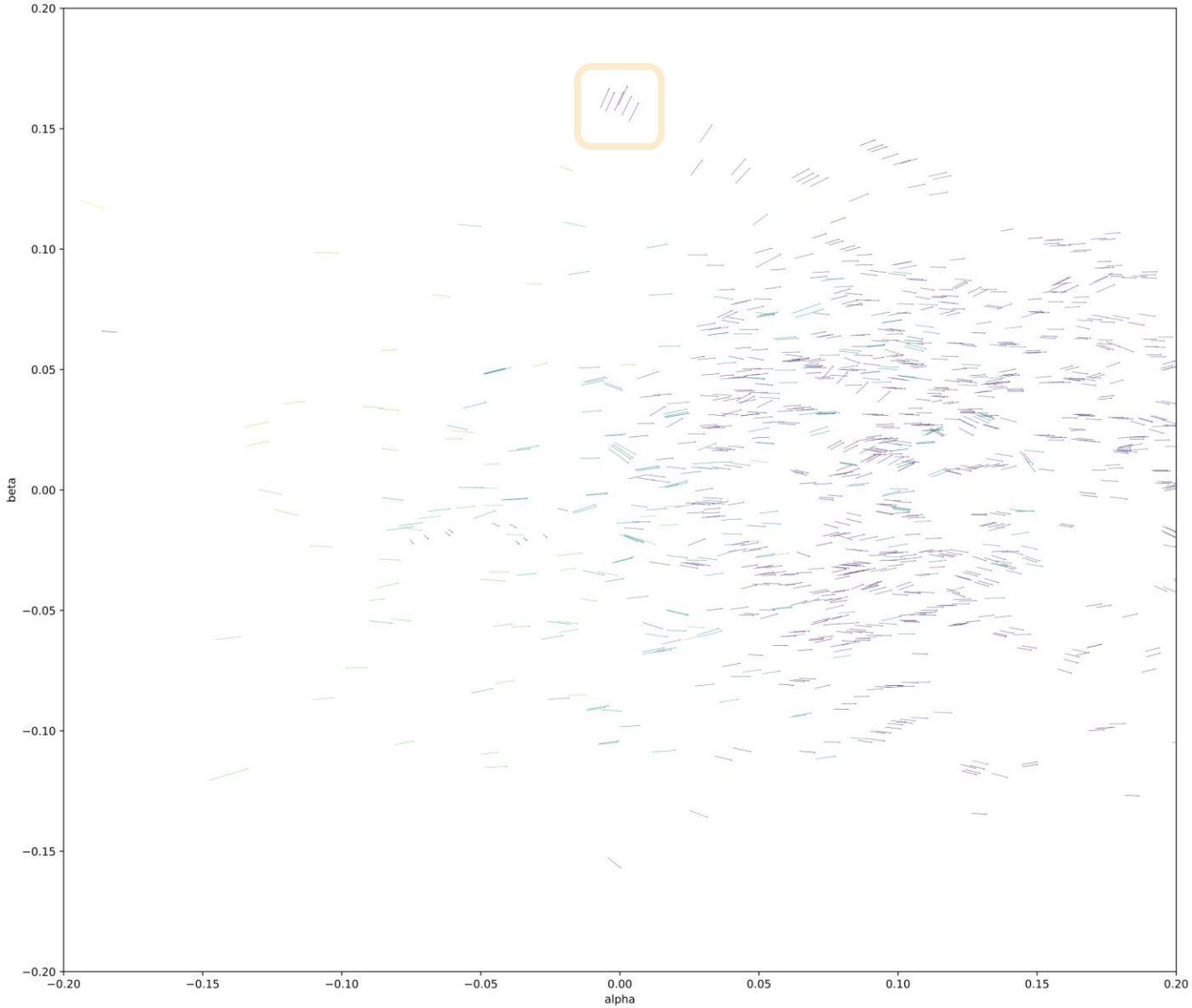
PROPOSED SOLUTION

- When we apply a heliocentric transformer, the trajectories become clear!
- Similar to applying the kernel trick in 3-dimensional space
- Can then set a 'focal distance' by inferring z-distance

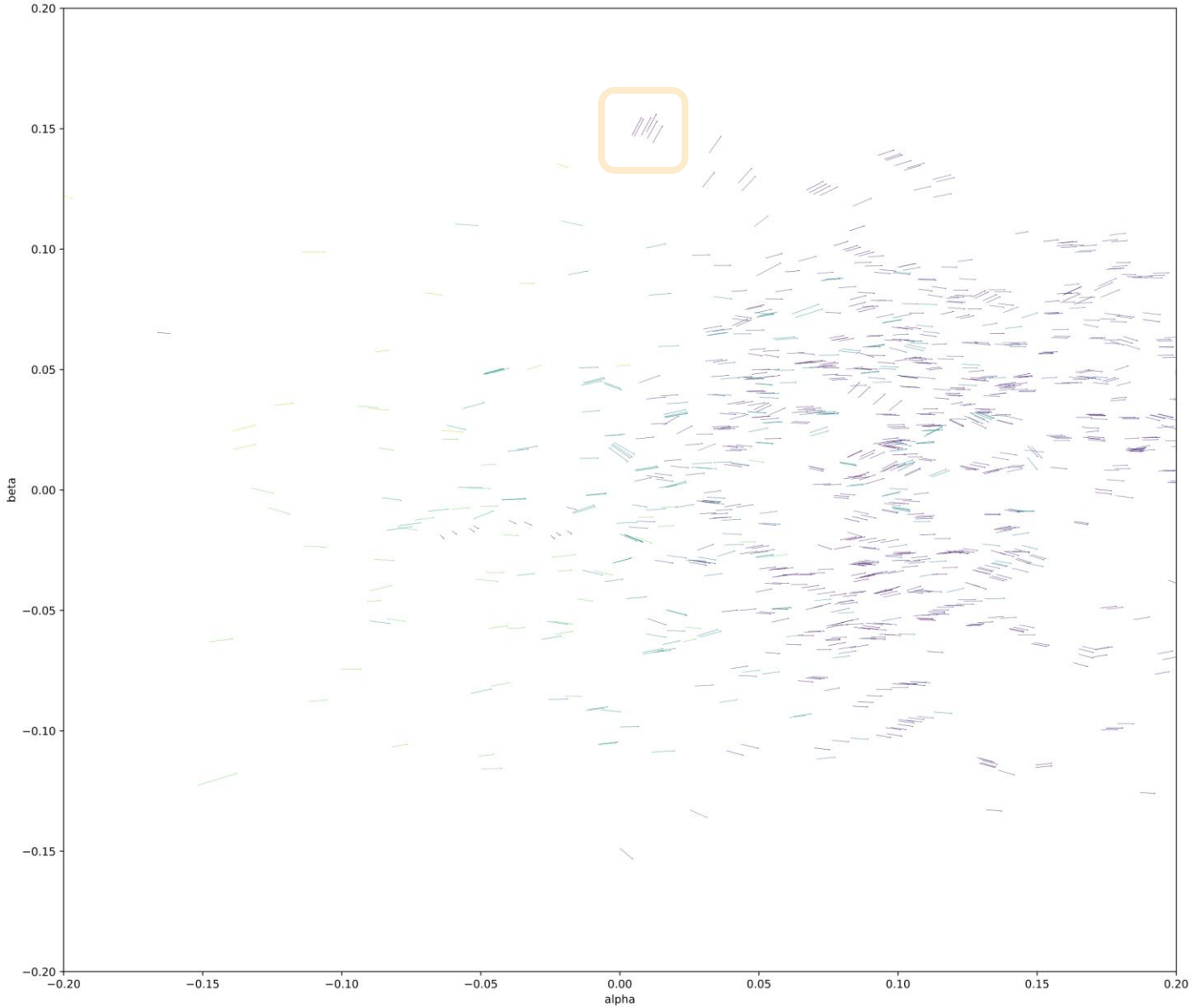
Partial asteroid tracklets (from the Sun's perspective)



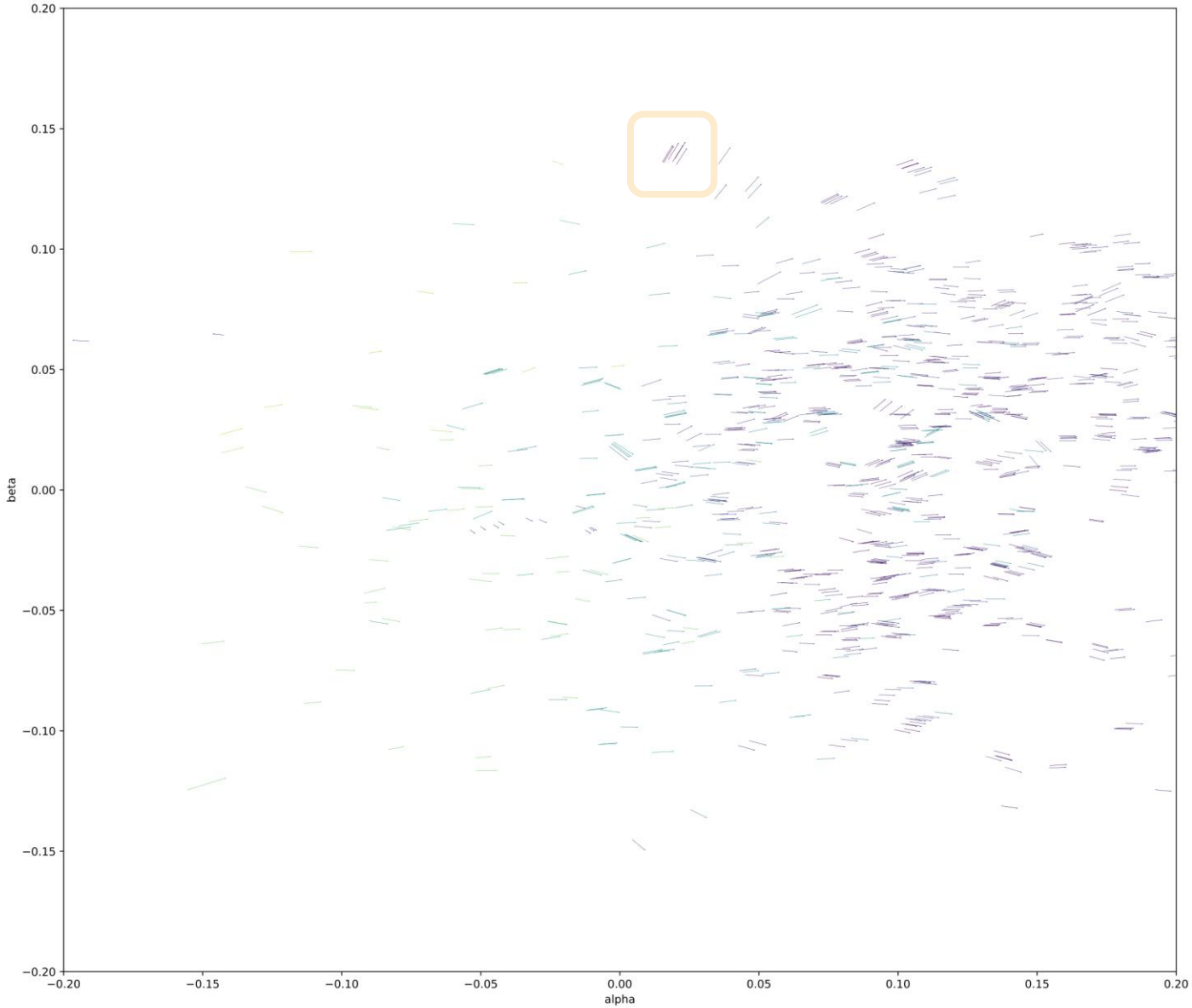
THE VIEW FROM THE SUN



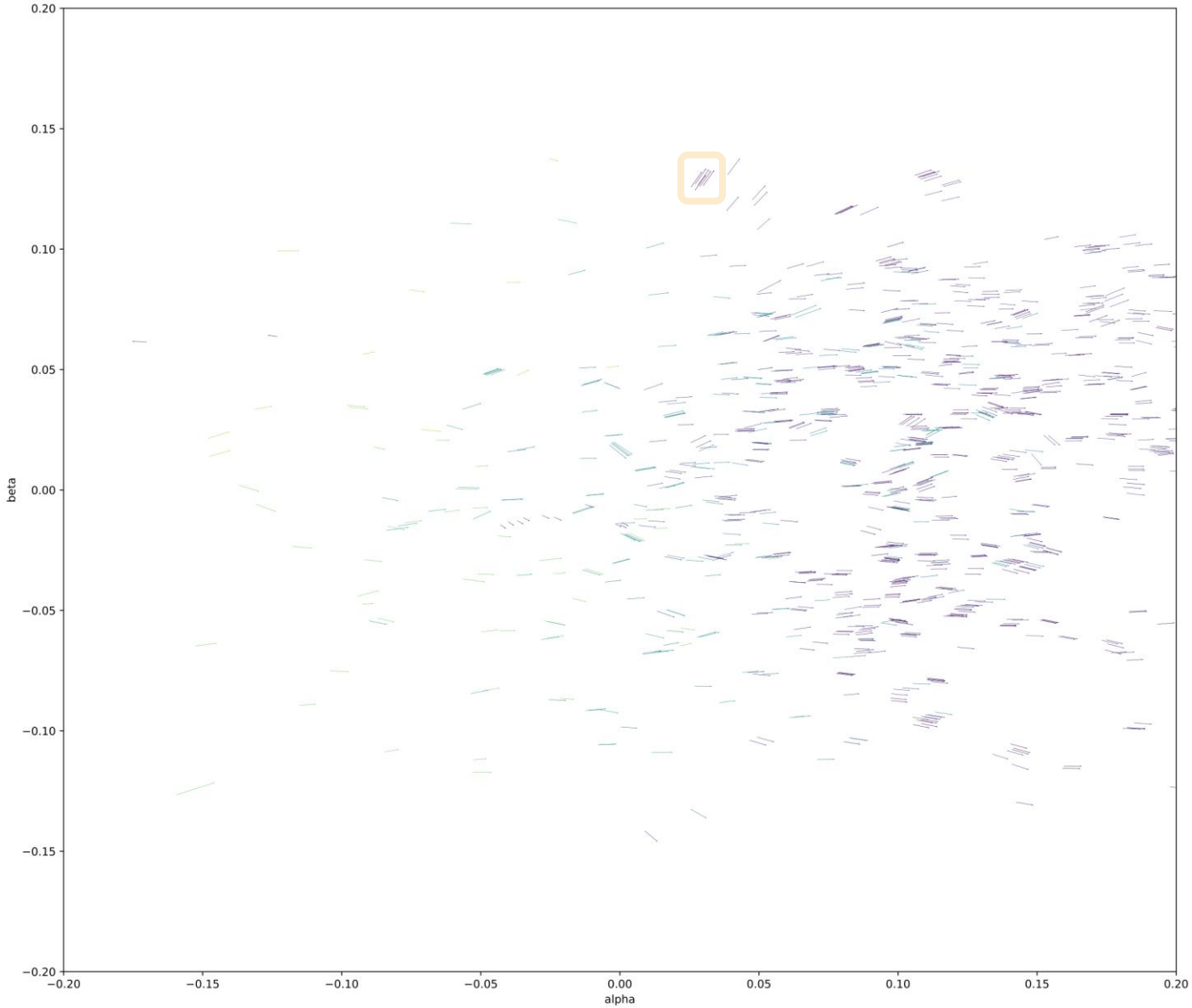
THE VIEW FROM THE SUN



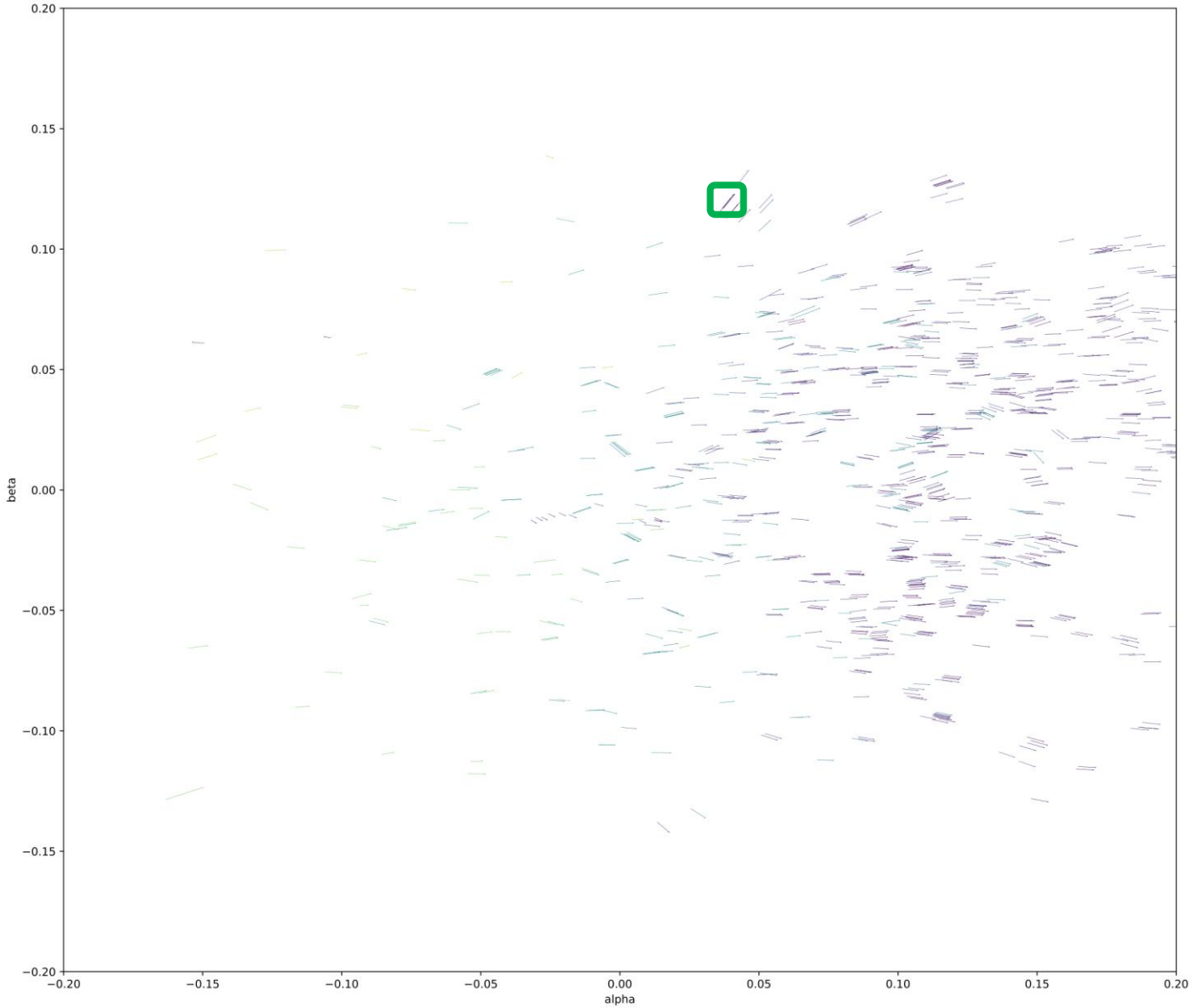
THE VIEW FROM THE SUN



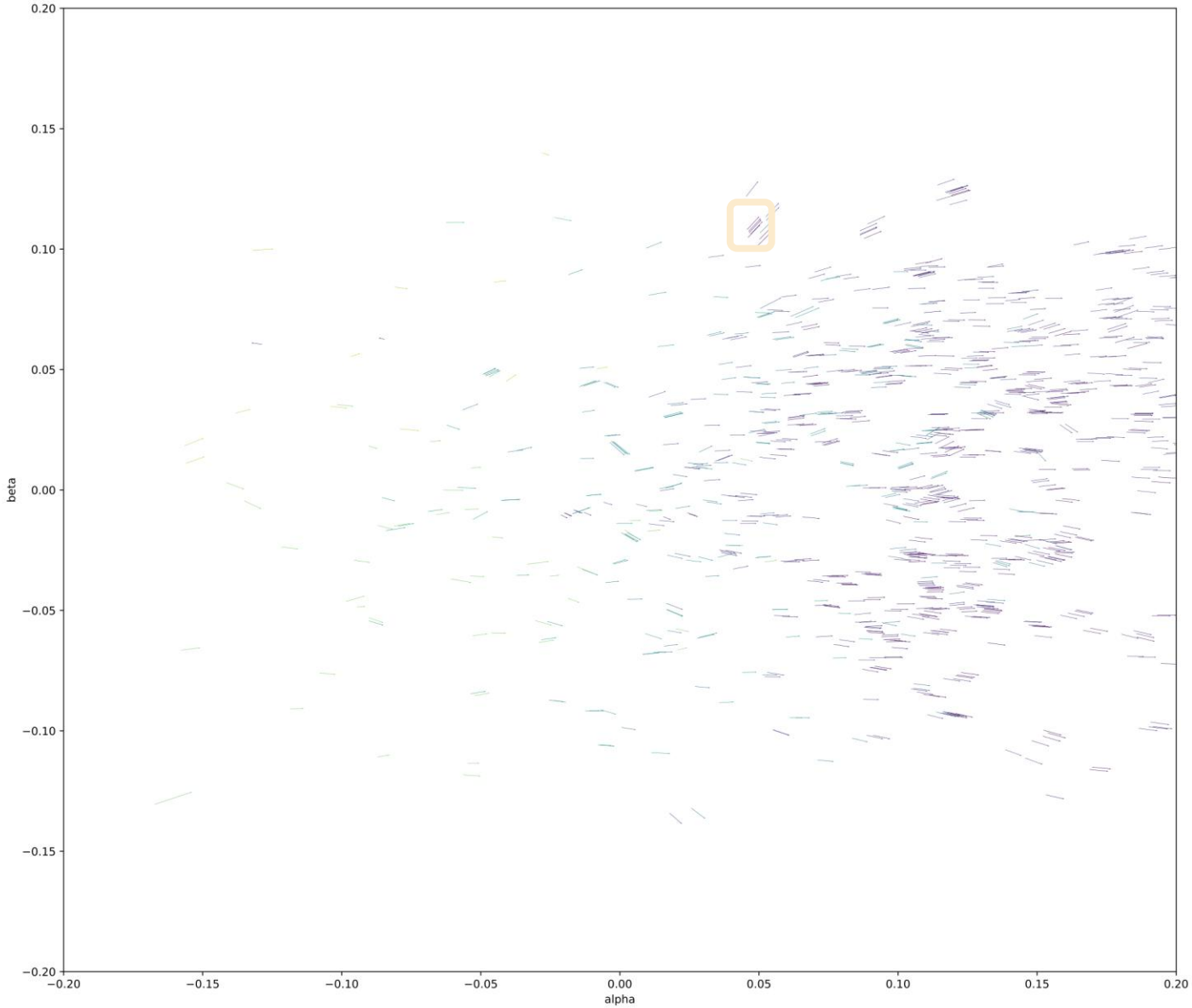
THE VIEW FROM THE SUN



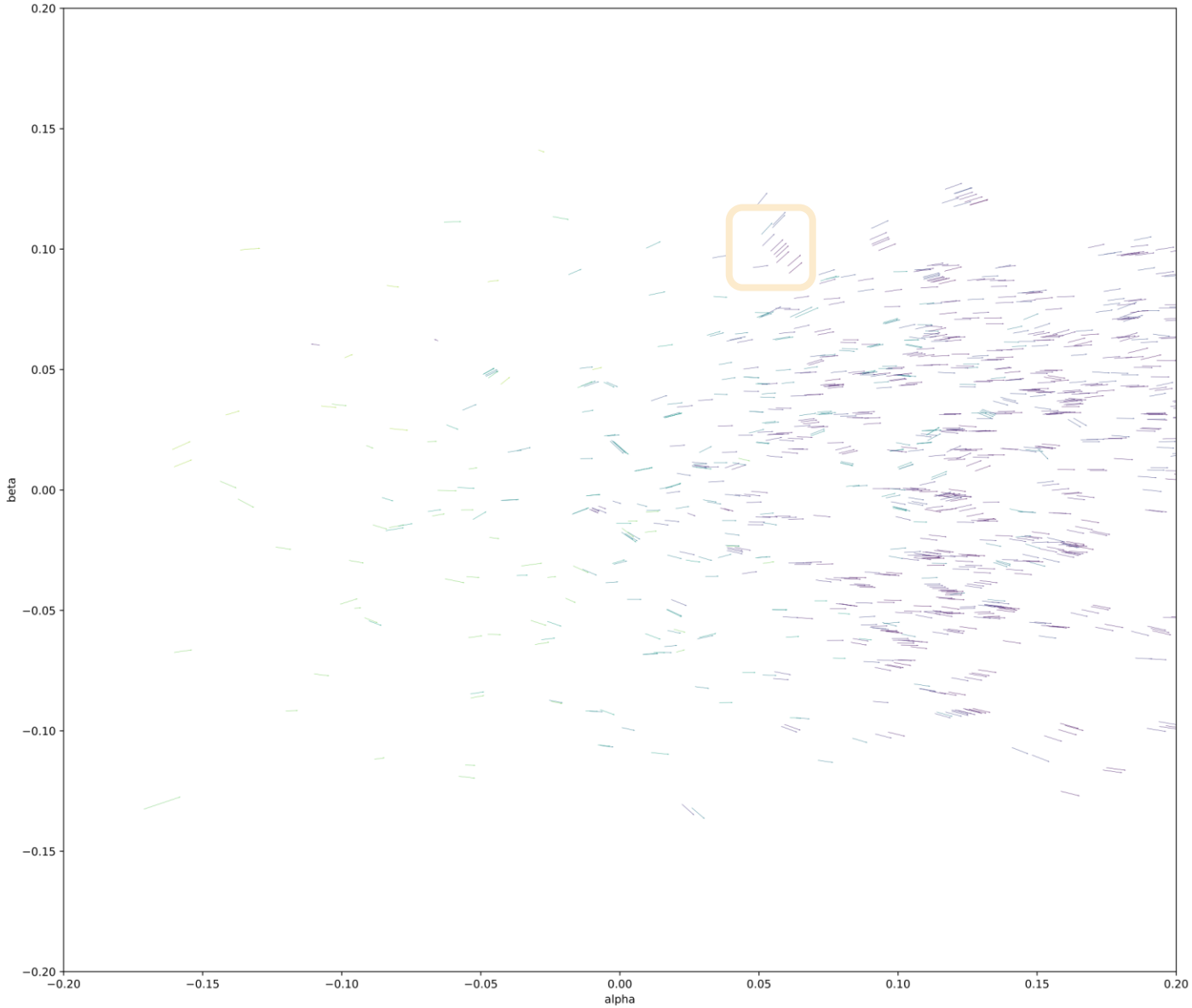
THE VIEW FROM THE SUN



THE VIEW FROM THE SUN



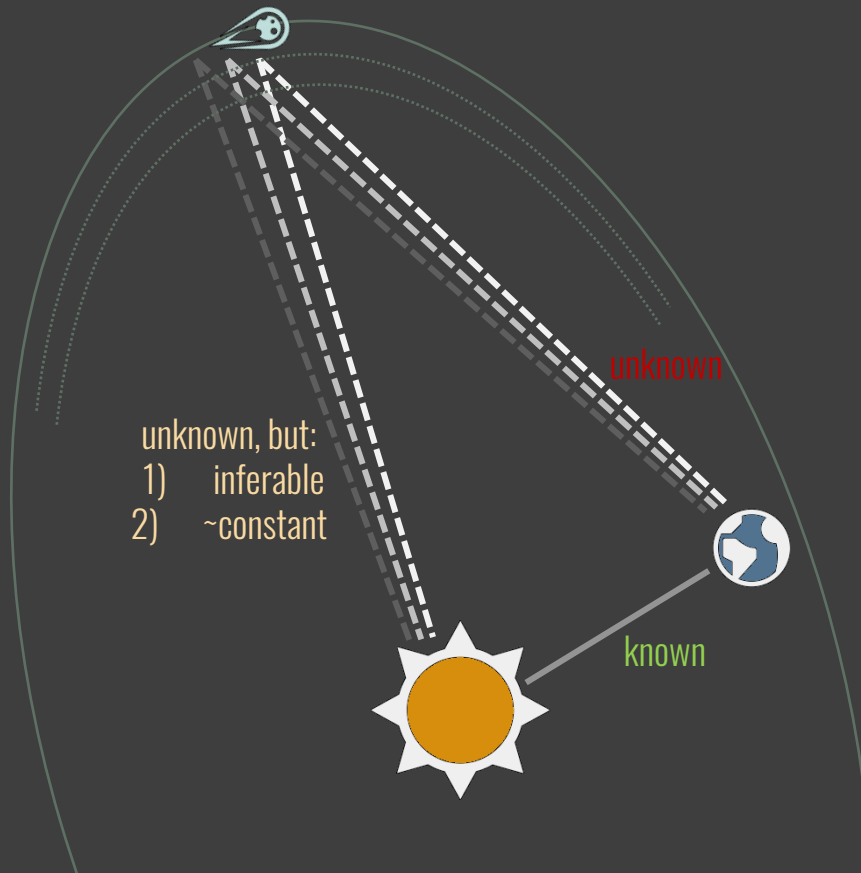
THE VIEW FROM THE SUN

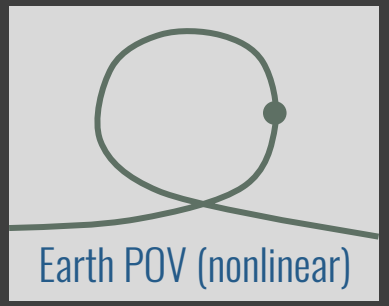


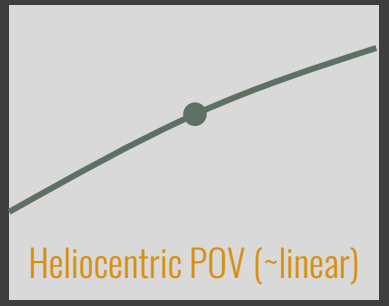
CORE INTUITION

WHY DOES HELIOCENTRICITY WORK?

- Heliocentric transform allows us to infer a constant distance from the sun
- Also helps us compare asteroid tracklets linearly, and in common time
- Enables clustering over small radii







*not to scale
(...obviously)

OUR METHODOLOGY

1. DIVIDE SKY INTO !!! PLANES AND 300 TIME SLICES

2. ASSERT THE (UNKNOWN) Z-DISTANCE FROM THE SUN

Running full sky in a single cluster is computationally intractable [O(n log n)]

OUR METHODOLOGY

1. DIVIDE SKY INTO !!! PLANES AND 300 TIME SLICES
2. ASSERT THE (UNKNOWN) Z-DISTANCE FROM THE SUN
3. ITERATE OVER Z-VELOCITIES, SOLVE OTHER MOTION PARAMS
4. CLUSTER BASED ON 6 PARAMS: X/Y/Z POSITION, X/Y/Z - VELOCITY

4 of our 6 motion parameters are known, and 1 is relatively constant

Model is linear in the parameters

OUR METHODOLOGY

TUNE CLUSTERING HYPERPARAMETERS OVER KNOWN TRAINING SET (THANKS TO MINOR PLANET CENTER!)

TRAIN CLUSTER RADIUS/WEIGHTING

1. DIVIDE SKY INTO !!! PLANES AND 300 TIME SLICES
2. ASSERT THE (UNKNOWN) Z-DISTANCE FROM THE SUN
3. ITERATE OVER Z-VELOCITIES, SOLVE OTHER X/Y MOTION PARAMS
4. CLUSTER BASED ON 6 PARAMS: X/Y/Z POSITION, X/Y/Z - VELOCITY

We tune two hyper-parameters:

1) OPTIMAL **CLUSTER RADIUS**

2) **WEIGHTING** OF VELOCITY VS. POSITION

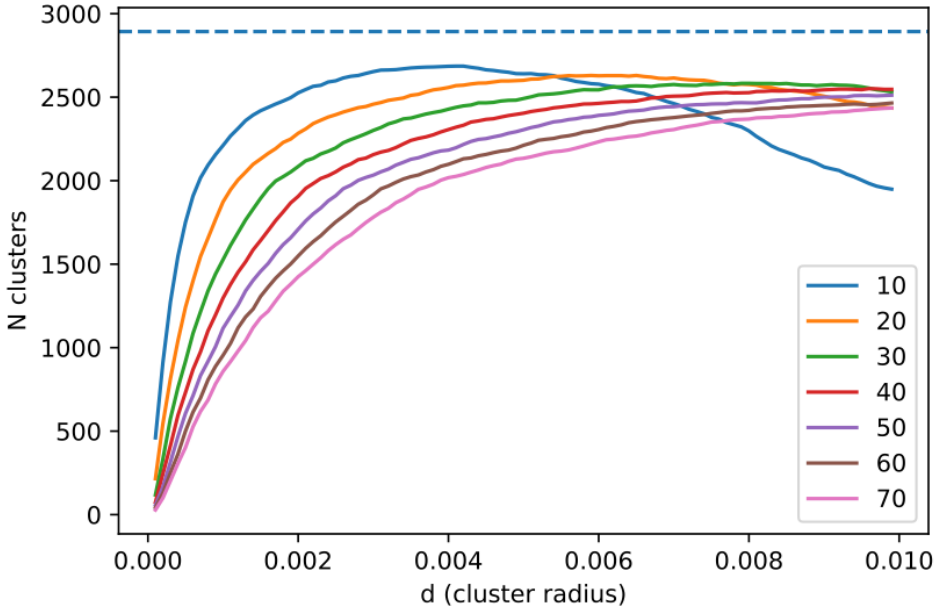
TUNING PERFORMANCE

OPTIMAL CLUSTER RADIUS: 0.005-0.010AU

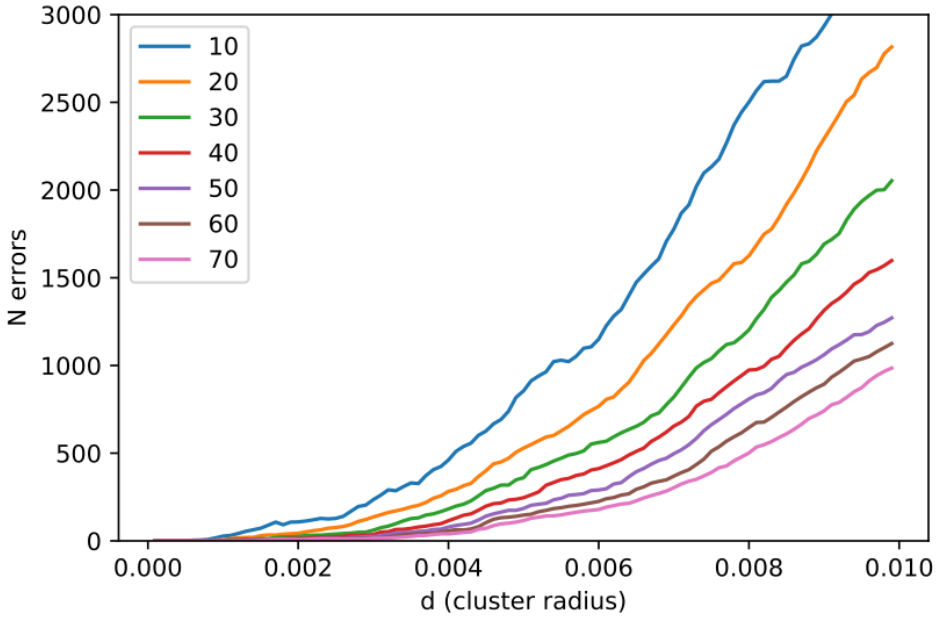
Selected KD-tree clustering for optimal speed (!!!)

Also had good results with agglomerative clustering

Total clusters detected vs. KD-tree cluster radius (training set)



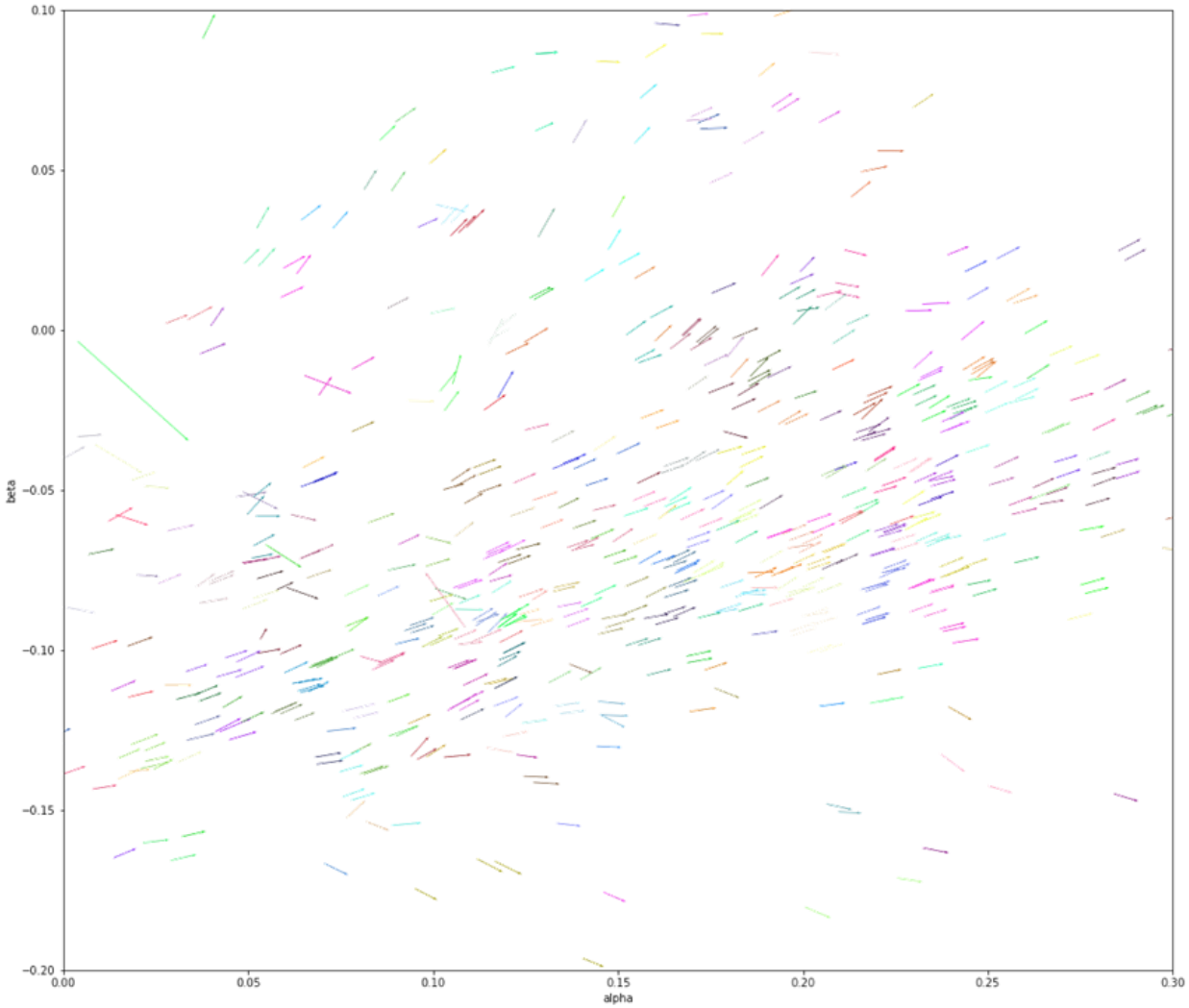
Error rate vs. KD-tree cluster radius (training set)



CLUSTERING RESULTS

90% OF ASTEROIDS DETECTED WITH ZERO ERROR!

Identified Asteroids (common colors = a single asteroid)



!!! COPY HERE?

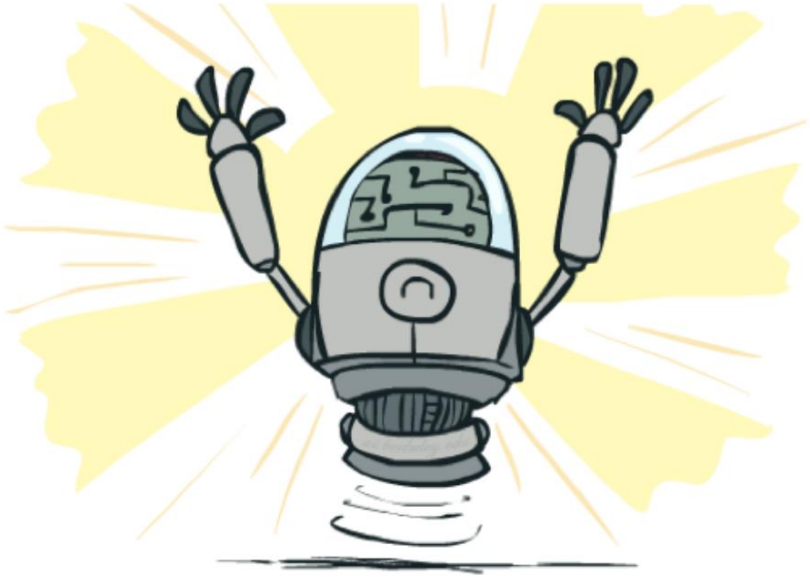
CONCLUSIONS

Conclusion: Successful tracklet matching using the heliocentric transform and KD-Trees

Also able to do so in just **O(n log n)** time, a scalable solution!!

Limitations: possible improvement in a more accurate clustering method

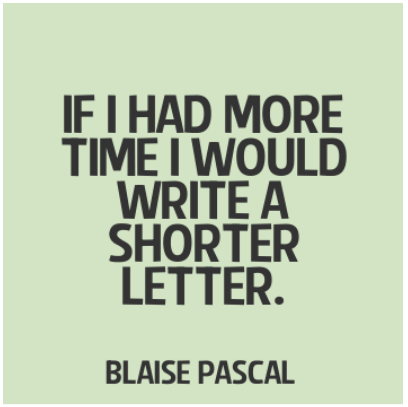
Next Steps: Run the algorithm over all 14m partial observations!



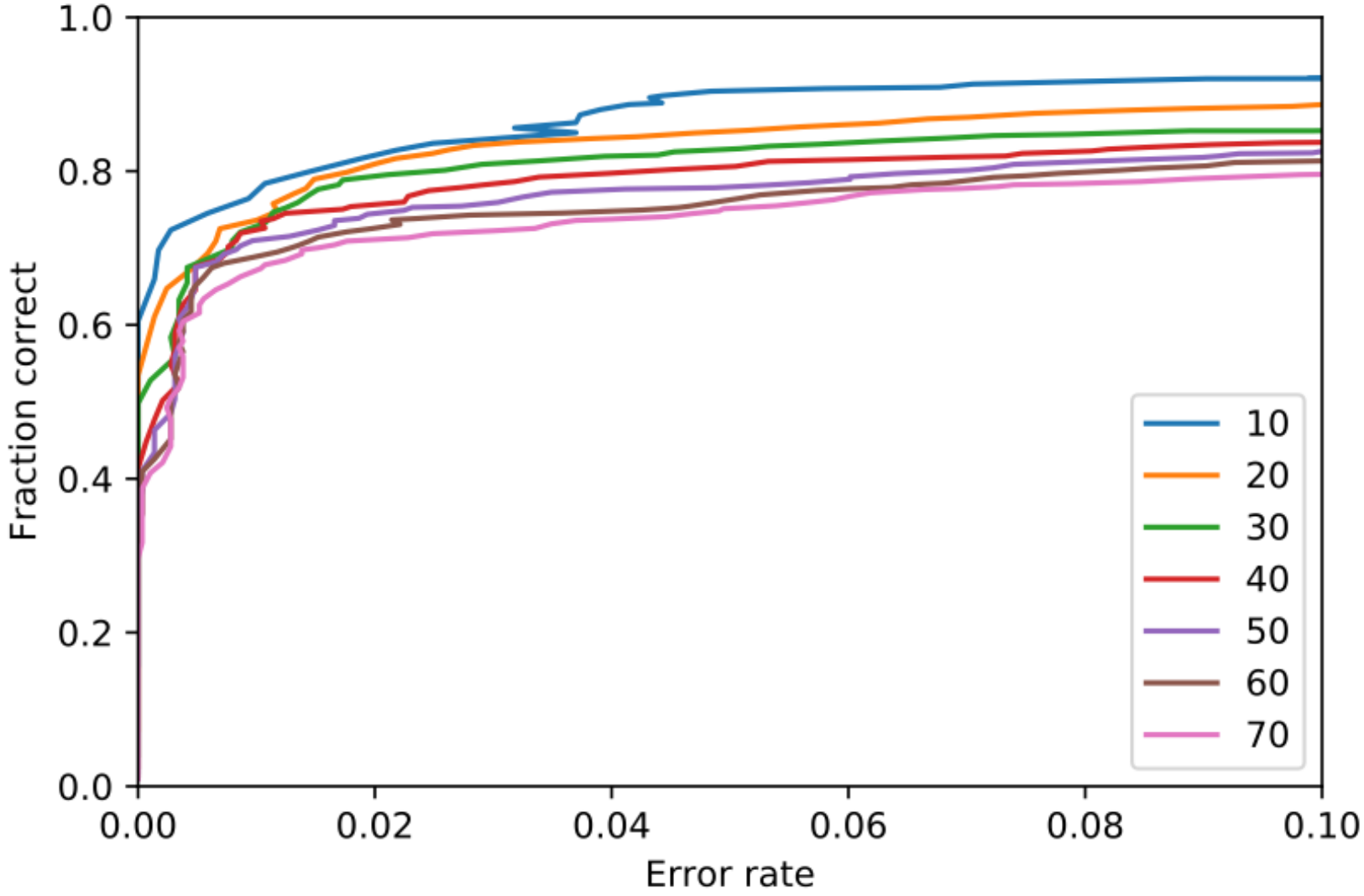
Thanks To: The Minor Planet Center for training data, Scott and Brian for project support!


00
00
00
1 COIN 1 PLAY
©1979 ATARI INC

APPENDIX/SUPPORTING SLIDES



Training AUC



Agglomerative Clustering Dendogram (Truncated)