

# Classification and Reconstruction from Single Lines

Donovan Webb

eBIC/University of Bath

March 23rd, 2020

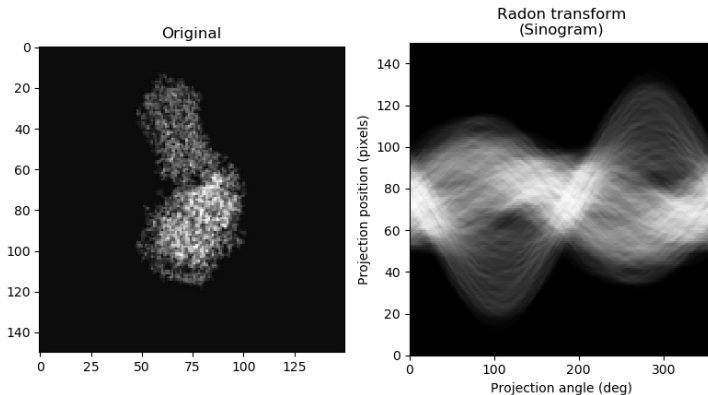


# Table of Contents

- 1 Single Lines
- 2 Finding Common Lines
- 3 Reconstruction
- 4 Clustering - Not done yet

# Single Lines

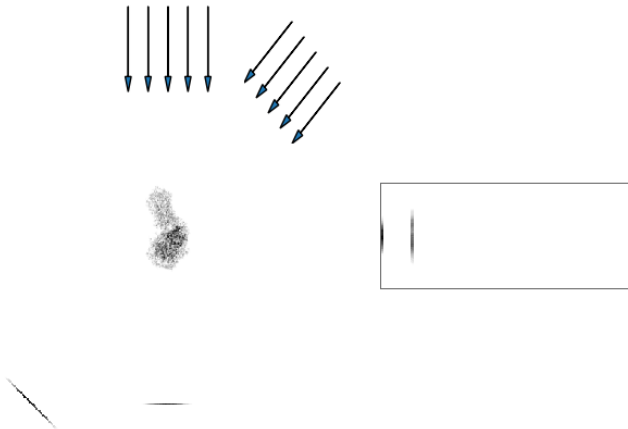
# The Radon Transform



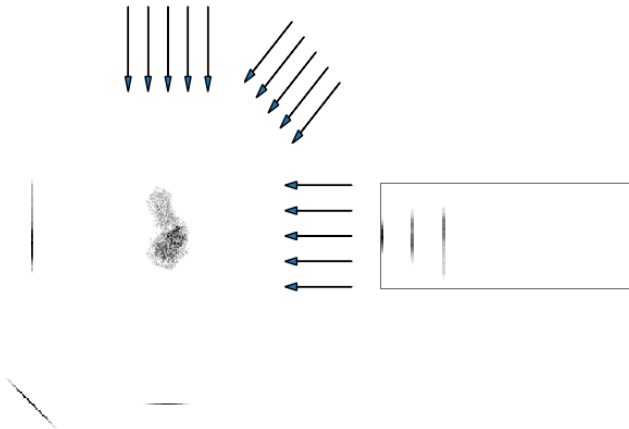
# The Radon Transform



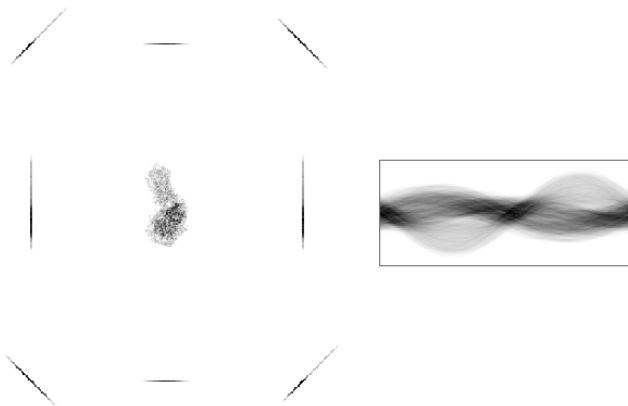
# The Radon Transform



# The Radon Transform



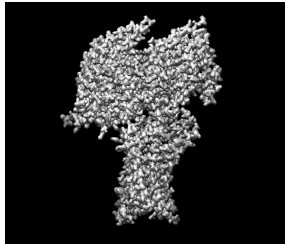
# The Radon Transform



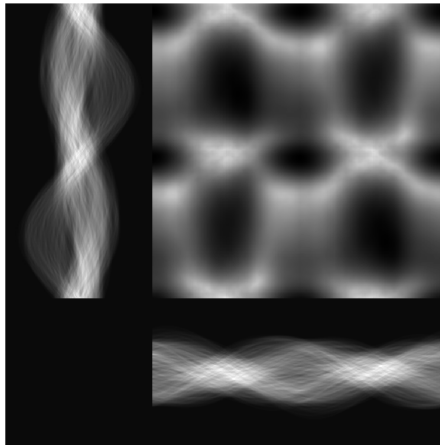


# Common Lines

**Two projections of the same 3D volume share at least one common line in the Radon transform**

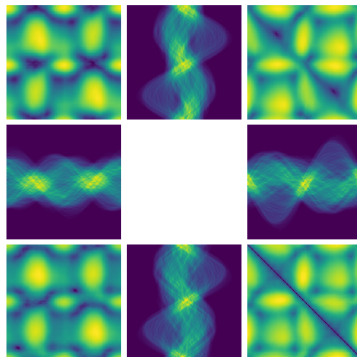


# Sinogram Cross Correlation



# Detour! Angular recovery from 3 Common lines

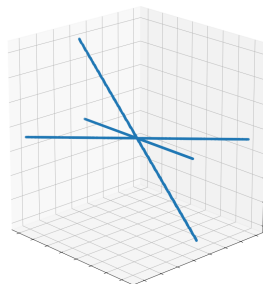
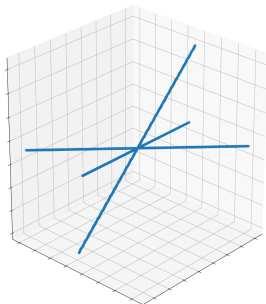
**Common line gives axis of rotation.** Three common lines gives 2 unique solutions for 3D orientation (One mirror of other)



*Angular Reconstitution: A Posteriori Assignment of Projection Directions for 3d Reconstruction. Van Heel 1987*

# Detour! Angular recovery from 3 Common lines

**Common line gives axis of rotation.** Three common lines gives 2 unique solutions for 3D orientation (One mirror of other)



*Angular Reconstitution: A Posteriori Assignment of Projection Directions for 3d Reconstruction. Van Heel 1987*

# Detour2! crYOLO

Particle picking improvements mean no longer have to rely on 2D classes to get sinograms from... Good centering, only particles being picked. Extensive use at eBIC will touch at the end...

## Finding Common Lines

# Raw single lines

Procedure for finding common lines. Get euclidian distances.  
Smallest is most similar

# Dimensional Reduction

Dimensional reduction can be used to lower influence of noise and try to find common features of data. Manifold learning found to improve accuracy of common lines being chosen although more computationally expensive. Plot 2D pca of sinogram and non-linear Iso map or lle



# Comparison between linear and non-linear dimensions

Graph of how lin vs non-lin change for different noise levels and different number of components

## Bonus slide: NN architecture for comparing single lines

Could we use NN to find filters? Number of attempts but nothing stable as yet. Probably leave out as unsuccessful and long to explain - could do a passing mention

# Reconstruction

# Voting

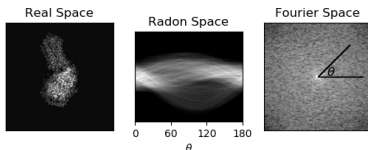
Take all lines pairs under distance  $X$ . This gives multiple possible axis between the same two projections. Weight these according to distance and number of times they appear. For each common line between two sinograms get more than 1 match so... vote which is best!

## Over determined system

3 lines is all good for getting angular information back c.f. prev slide. any more than 3 and we can get contradicting equations - system is overdetermined. we have  $N^2$  linear equations and only  $6N$  variables (do proper maths) Machine learning 3! how to fit. Least squares optimization not good as we have large number of misidentified lines and nonconvex! we also need to get proper rotation matrices at the end. This constraint leads to collapse to trivial solution 000. Need other approach.

# Eigenvector Relaxation

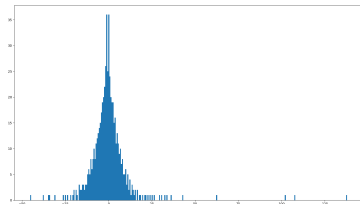
**Aim: Given all common lines  $c$  for projections  $P$ , assign Rotation matrices  $R$  for each  $P$  to give greatest consensus volume.**



$$c_{ij} = (\cos(\theta_{ij}), \sin(\theta_{ij}), 0), \quad c_{ji} = (\cos(\theta_{ji}), \sin(\theta_{ji}), 0) \quad (1)$$

$$\max \sum_{i \neq j} R_i c_{ij} \cdot R_j c_{ji} \quad (2)$$

Maths\*! Make large  $(2N \times 2N)$  symmetric matrix  $S$ . Can recover  $R$  for each  $P$  from top 3 eigenvectors of  $S$  that maximise (2)!



*\*Three Dimensional Structure Determination from Common Lines in Cryo-EM by Eigenvectors and Semidefinite Programming. A.Singer and Y.Shkolnisky*

# Models!

Lots of models!

# Computational efficiency

only calculating 95% of correct lines make others random. 6x speed increase. Parrallelisation needed! Show that adding random lines does not affect recon too much..



## Priming matrix with tilt data! - Not done yet

With tilt data we know the tilt axis (and angle but this is not important) Can calculate what common single line would be for this data would be. can input straight into matrix

Clustering - Not done yet

# Can heterogeneity be sorted by looking at common lines?

A

- Lorem Ipsum
- Dolor est
- Example incoming
- Example arrived

B

- Sometimes its hard to make up random content
- Othertimes not
- Filler words
- And filler phrases that are ever so slightly longer

# Can heterogeneity be sorted by looking at common lines?

## A

- Lorem Ipsum
- Dolor est
- Example incoming
- Example arrived

## B

- Sometimes its hard to make up random content
- Othertimes not
- Filler words
- And filler phrases that are ever so slightly longer