Robust Statistical Methods for the Rotation Group

Bryan Stanfill

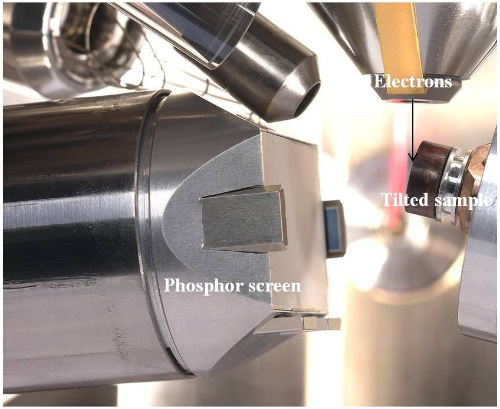
17th International Conference on Information Fusion

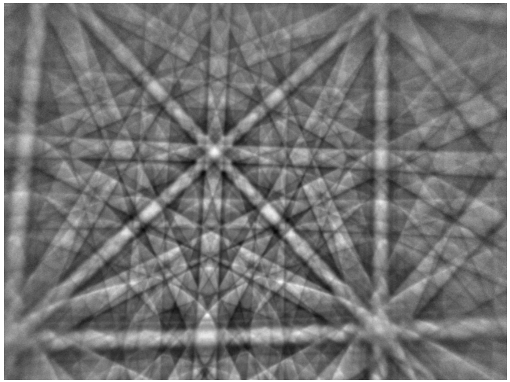
COMPUTATIONAL INFORMATICS
www.csiro.au

Electron Backscatter Diffraction[a]

[a]images from www.EBSD.com

- Electron backscatter diffraction (EBSD) is used to learn about the microstructure of crystalline materials
- Phase discrimination, texture analysis, crystal orientation mapping...

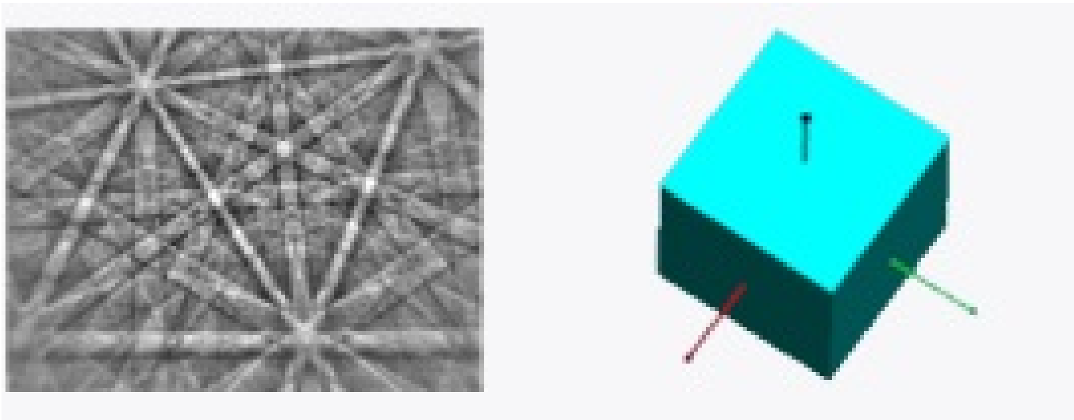




Crystal Orientation[a]

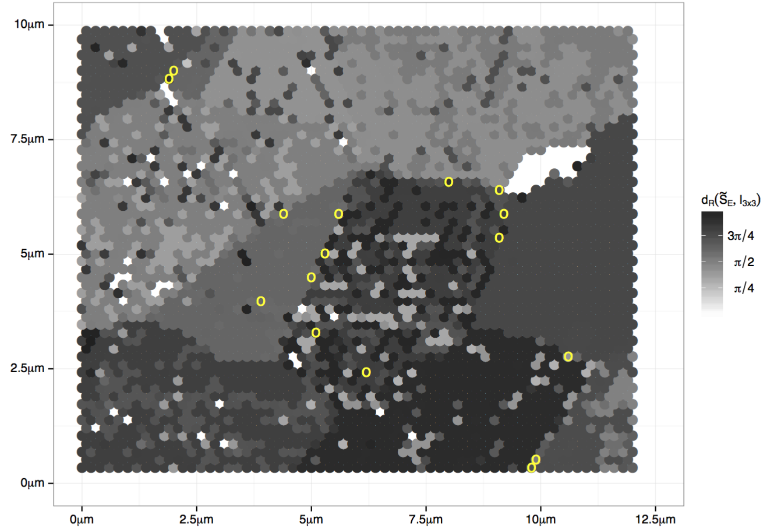
[a]images from www.EBSD.com

- Each image is translated into an orientation
- Orientations are expressed as 3×3 orthogonal matrices with determinant $+1$
- Crystal orientations are crucial in texture analysis and grain boundary identification



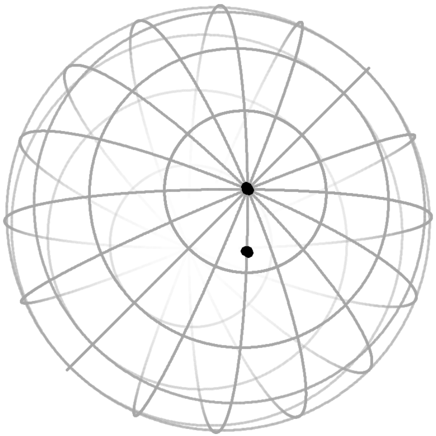
Example Data

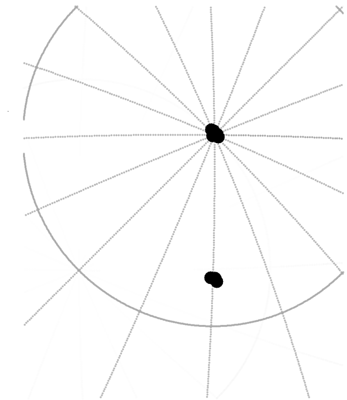
- We consider EBSD data obtained by scanning a nickel surface 14 times
- Interest is in the orientation of cubic crystals on the metal surface at a each location
- How should multiple scans be summarized?



Extreme observations

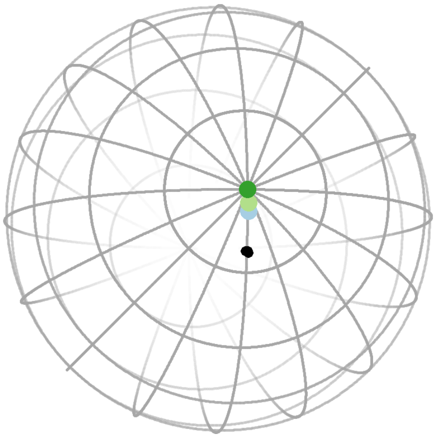
- When multiple scans are available they are combined to form one coherent grain map
- For locations on the boundary of two grains, the measurements can be mix of the two grains
- In that case the mean doesn't represent either grain but other estimators can

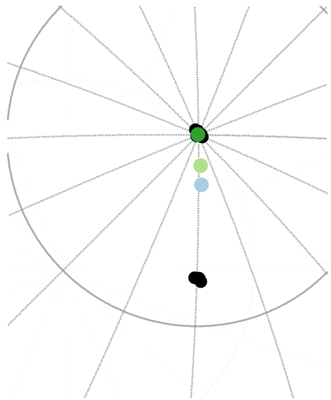




Extreme observations

- When multiple scans are available they are combined to form one coherent grain map
- For locations on the boundary of two grains, the measurements can be mix of the two grains
- In that case the mean doesn't represent either grain but other estimators can





In This Talk

1. Introduce a measure of discord for the rotation group
2. Propose a new estimator based on that measure
3. Introduce the influence functions for the rotation group to better understand estimator behavior
   - Establishing a hierarchy of robust estimators
   - Understand when to use which estimators

Matrix Representation of Rotations

- $SO(3)$ - collection of all 3×3 orthogonal matrices $\boldsymbol{R}$ with $\det(\boldsymbol{R}) = 1$
- $\boldsymbol{R} \in SO(3)$ is associated with $\boldsymbol{W} = (w_1, w_2, w_3)^\top \in \mathbb{R}^3$

$$\boldsymbol{R} = \exp[\boldsymbol{\Phi}(\boldsymbol{W})] = \cos(r)\boldsymbol{I} + \sin(r)\boldsymbol{\Phi}(\boldsymbol{U}) + (1 - \cos r)\boldsymbol{U}\boldsymbol{U}^\top$$

where

- $\boldsymbol{\Phi}(\boldsymbol{W}) = \begin{bmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{bmatrix}$
- $r = \|\boldsymbol{W}\|$ and $\boldsymbol{U} = \boldsymbol{W}/\|\boldsymbol{W}\| \in \mathbb{R}^3$
- r and $\boldsymbol{U}$ termed misorientation angle and axis, respectively

Statistical Treatment of $SO(3)$

- $\boldsymbol{R}_1, \ldots, \boldsymbol{R}_n \in SO(3)$ random sample from

$$\boldsymbol{R}_i = \boldsymbol{S}\boldsymbol{E}_i,$$

  the $SO(3)$ analog to $y_i = \mu + \epsilon_i$ where

  - $\boldsymbol{S} \in SO(3)$ - parameter measuring central tendency
  - $\boldsymbol{E}_1, \ldots, \boldsymbol{E}_n \in SO(3)$ i.i.d. random rotations each with r_i, $\boldsymbol{U}_i$
  - $\boldsymbol{U}_i$ uniformly distributed on unit sphere
  - r_i distributed symmetrically about 0 on the interval $[-\pi, \pi)$
  - r_i and $\boldsymbol{U}_i$ are independent for all i

M-Estimators in $SO(3)$

- M-estimators for the central orientation $\boldsymbol{S} \in SO(3)$ are of the form

$$\widehat{\boldsymbol{S}} = \underset{\boldsymbol{S} \in SO(3)}{\arg\min} \sum_{i=1}^{n} \rho(\boldsymbol{R}_i, \boldsymbol{S})$$

- Three common choices of $\rho(\boldsymbol{R}_i, \boldsymbol{S})$
  - The projected mean $\rho(\boldsymbol{R}_i, \boldsymbol{S}) = \|\boldsymbol{R}_i - \boldsymbol{S}\|_F^2 \propto -\text{tr}(\boldsymbol{S}^\top \boldsymbol{R}_i)$
  - The projected median $\rho(\boldsymbol{R}_i, \boldsymbol{S}) = \|\boldsymbol{R}_i - \boldsymbol{S}\|_F$
  - The geometric median $\rho(\boldsymbol{R}_i, \boldsymbol{S}) = \frac{1}{\sqrt{2}}||\text{Log}(\boldsymbol{R}_i^\top \boldsymbol{S})||_F$

M-Estimators in $SO(3)$

- M-estimators for the central orientation $\boldsymbol{S} \in SO(3)$ are of the form

$$\widehat{\boldsymbol{S}} = \underset{\boldsymbol{S} \in SO(3)}{\arg\min} \sum_{i=1}^{n} \rho(\boldsymbol{R}_i, \boldsymbol{S})$$

- Three common choices of $\rho(\boldsymbol{R}_i, \boldsymbol{S})$
  - The projected mean $\rho(\boldsymbol{R}_i, \boldsymbol{S}) = \|\boldsymbol{R}_i - \boldsymbol{S}\|_F^2 \propto -\text{tr}(\boldsymbol{S}^\top \boldsymbol{R}_i)$
  - The projected median $\rho(\boldsymbol{R}_i, \boldsymbol{S}) = \|\boldsymbol{R}_i - \boldsymbol{S}\|_F$
  - The geometric median $\rho(\boldsymbol{R}_i, \boldsymbol{S}) = \frac{1}{\sqrt{2}} ||\text{Log}(\boldsymbol{R}_i^\top \boldsymbol{S})||_F$

Efficient and Robust Estimator

- We propose a weighted mean that balances robustness and efficiency

$$\operatorname*{arg\,min}_{\boldsymbol{S}\in SO(3)} \sum_{i=1}^{n} w_i \|\boldsymbol{R}_i - \boldsymbol{S}\|_F^2 = \operatorname*{arg\,max}_{\boldsymbol{S}\in SO(3)} \mathrm{tr}(\boldsymbol{S}^\top \overline{\boldsymbol{R}}_W)$$

where

- $\overline{\boldsymbol{R}}_W = \sum_{i=1}^{n} w_i \boldsymbol{R}_i$
- $w_i^{-1} \propto \sqrt{H_i}$ and

$$H_i = \frac{\sum_{i=1}^{n} \|\boldsymbol{R}_i - \widehat{\boldsymbol{S}}_E\|_F^2 - \sum_{j\neq i} \|\boldsymbol{R}_j - \widehat{\boldsymbol{S}}_E^{(i)}\|_F^2}{\sum_{j\neq i} \|\boldsymbol{R}_j - \widehat{\boldsymbol{S}}_E^{(i)}\|_F^2/(n-2)}$$

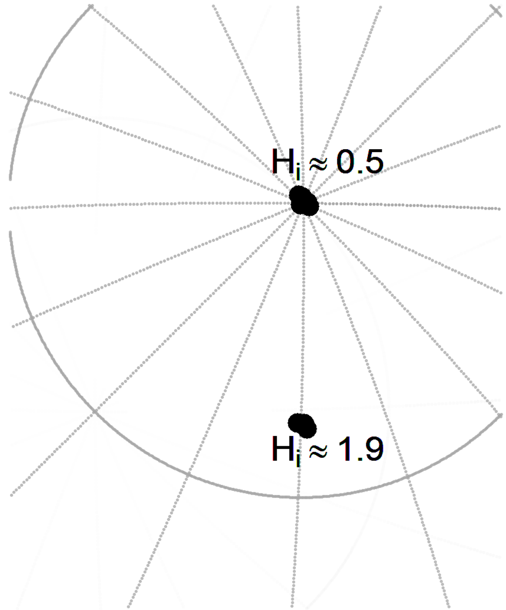
- $\widehat{\boldsymbol{S}}_E^{(i)}$ is the projected mean for the sample after omitting $\boldsymbol{R}_i$

More on H_i

- H_i is a discord measure originally proposed by Best and Fisher (1986) for polar data on the sphere

- If $\boldsymbol{E}_i \sim$ von Mises-Fisher then $H_i \sim F_{1,n-2}$

- If $\boldsymbol{E}_i \sim$ Cayley or matrix Fisher then $H_i \sim F_{3,3(n-2)}$

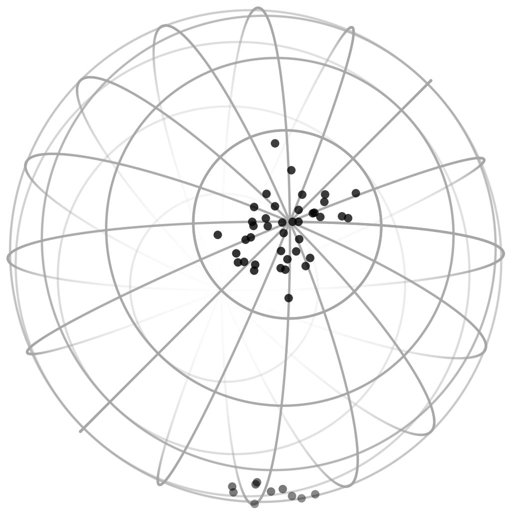


Simulation Study Design

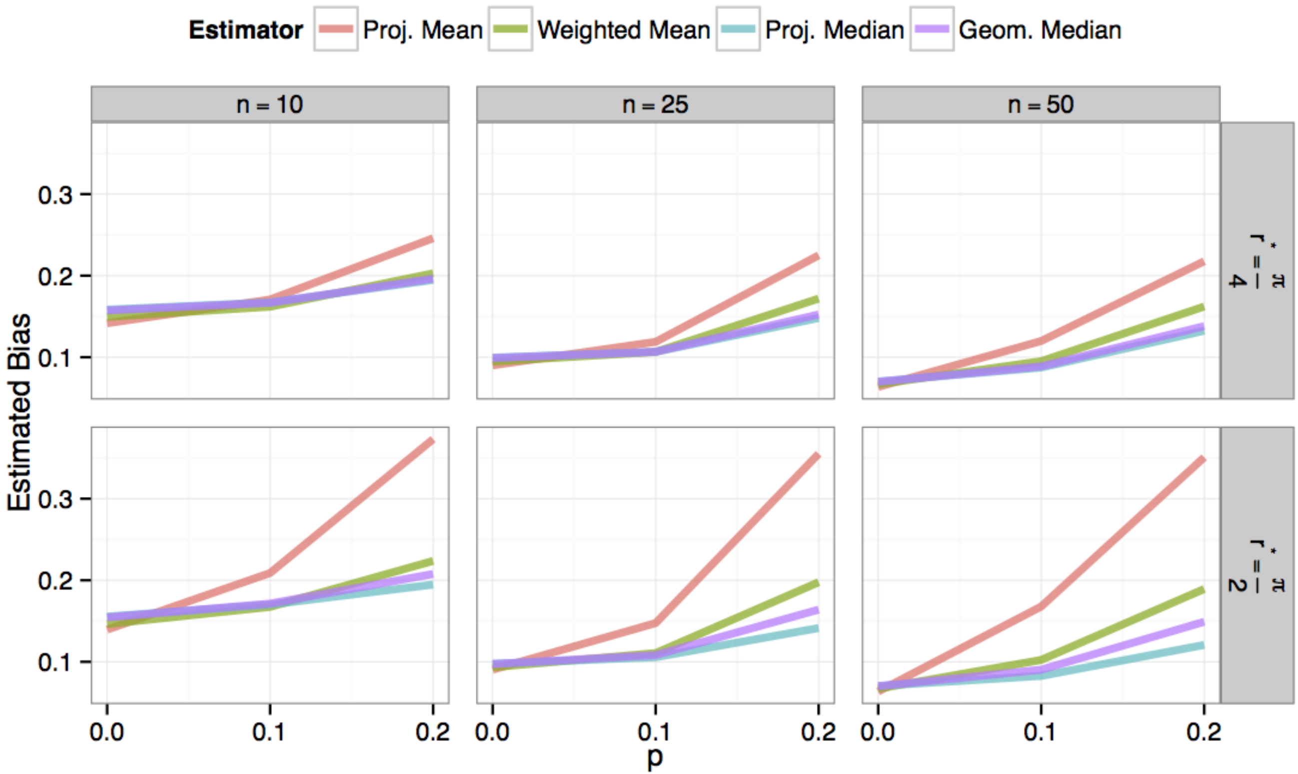
- Samples $\boldsymbol{R}_1, \ldots, \boldsymbol{R}_n$ were generated from the model

$$\boldsymbol{R}_i = \boldsymbol{S}\boldsymbol{E}_i I(i \in \mathcal{G}_1) + \boldsymbol{S}^*\boldsymbol{E}_i[1 - I(i \in \mathcal{G}_1)]$$

  - $I(i \in \mathcal{G}_1) = \begin{cases} 1 & \text{if } i \in \mathcal{G}_1 \\ 0 & \text{otherwise} \end{cases}$
  - $\boldsymbol{E}_i \sim$Cayley$(\boldsymbol{I}_{3\times 3}, \kappa = 50)$
  - $P(i \notin \mathcal{G}_1) = 0, 0.1, 0.2$
  - $n = 10, 25$ and 50
  - $\boldsymbol{S}^* = \exp\left[\boldsymbol{\Phi}\left([r^*, 0, 0]^\top\right)\right]$
  - $r^* \in \{\pi/4, \pi/2\}$

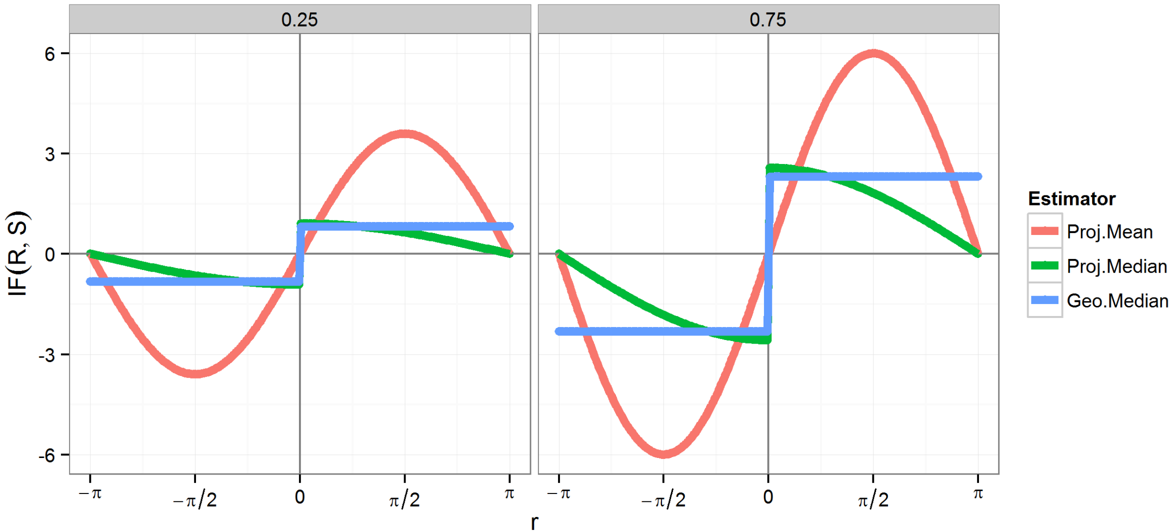


Simulation Study Results



Influence Functions

- Influence functions (IF) have been derived to explain estimator behavior



Discussion and Future Research

- The weighted mean has promising empirical properties but no theoretical results are available
- Other robustified L_2 estimators are of interest: winsorized mean, trimmed mean, Huber estimators
- IFs are one way to compare extrinsic and intrinsic approaches to $SO(3)$ data analysis
- We have approximated the IF for the geometric median but a closed form solution is currently unavailable

Thank You

CSIRO Computational Informatics

Bryan Stanfill

t +61 7 3833 5727

e Bryan.Stanfill@csiro.au

w Computational Informatics website