# A Gradient Field Approach to Modelling Fibre-Generated Spatial Point Processes

Bryony J. Hill and Wilfrid S. Kendall and Elke Thönnes
B.J.Hill@warwick.ac.uk, W.S.Kendall@warwick.ac.uk, E.Thonnes@warwick.ac.uk

Dublin ISI, STS031
Modeling and Analysis of Complex Data Objects
23 August 2011

Warwick Statistics





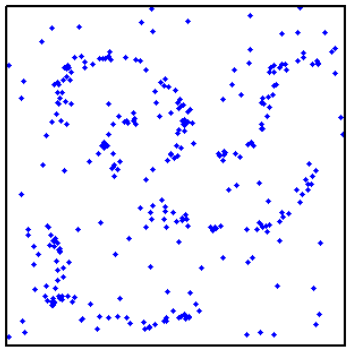
## Introduction

- How to infer unobserved curvilinear structure?
- Specifically,
  - Given observed pattern of points in plane or in space,
  - draw reasonable conclusions about unobserved curvilinear fibres along which points are supposed to cluster:
    - Number of fibres,
    - Total fibre length,
    - Signal points *versus* noise points,
    - Classification of points by fibre,
    - Orientation distributions of fibres,
    - ...

Warwick Statistics

# Illustrative planar point pattern examples

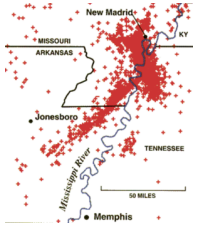
5: Fibre-generated Poisson Process

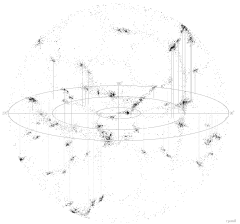






# Three typical application contexts:



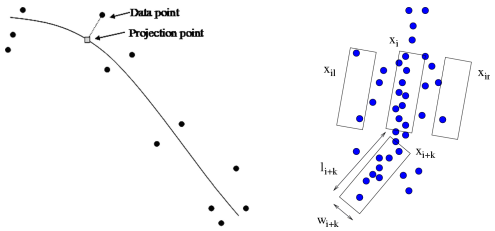
1. Fingerprint sweat pores
   Extracted from fingerprint a002-5 from NIST Special database 30 (Watson 2001).
2. Earthquake epicentres

   Epicentres in New Madrid region,
   taken from CERI (Center for Earthquake Research and Information).
3. Universe within 500 Mly

   Image: Richard Powell (`atlasoftheuniverse.com/nearsc.html`:
   Creative Commons Attribution-ShareAlike 2.5 License).

# Previous approaches

These include:







- Principal curves (Stanford and Raftery 2000)
- Candy, Bisous models (Stoica, Martínez, and Saar 2007)
- also, gradients of density estimates (Genovese, Perone-Pacifico, Verdinelli, and Wasserman 2009)

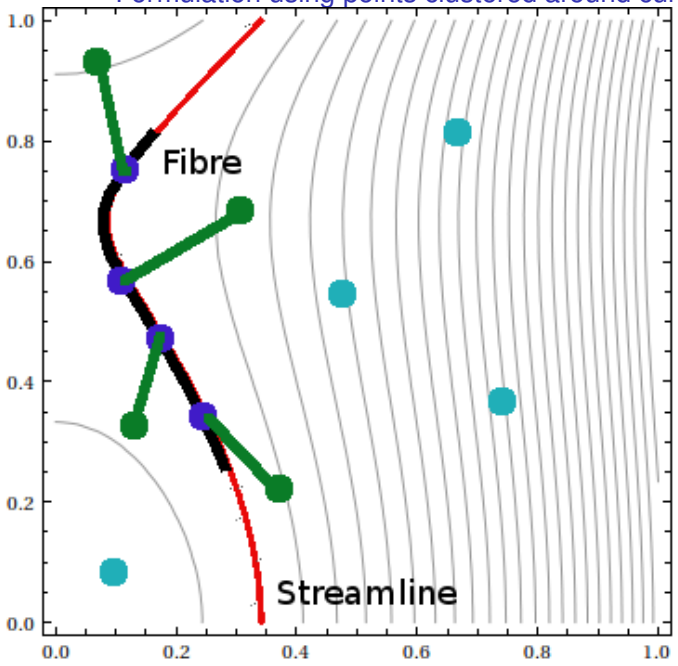
Warwick Statistics





# Our statistical model (I)

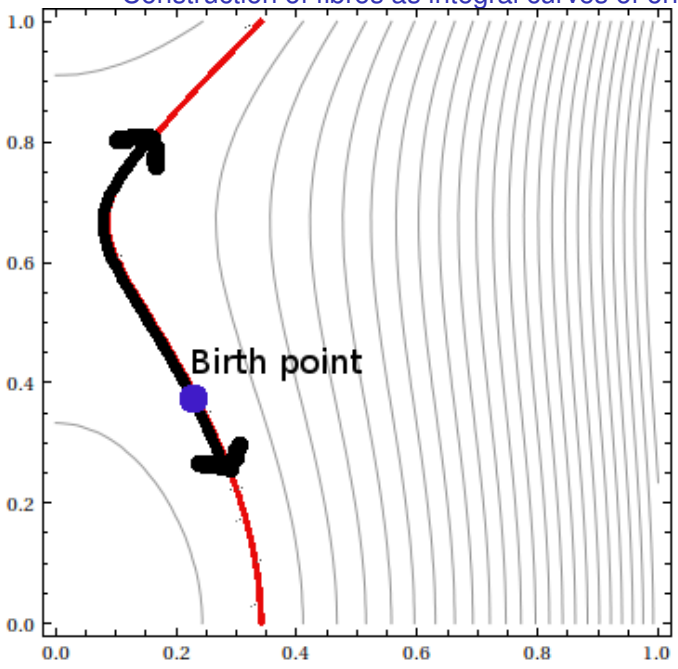
Formulation using points clustered around curvilinear fibres




Warwick Statistics

# Our statistical model (II)

## Construction of fibres as integral curves of orientation field



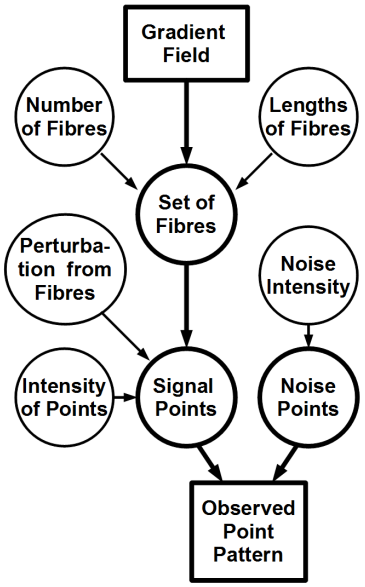

Warwick Statistics

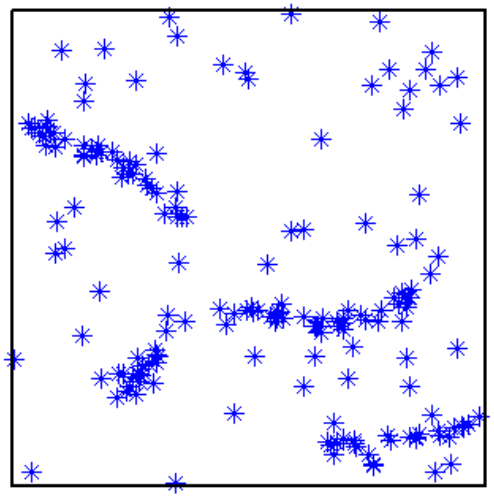




# Our statistical model (III)

## Building up a (simplified) DAG

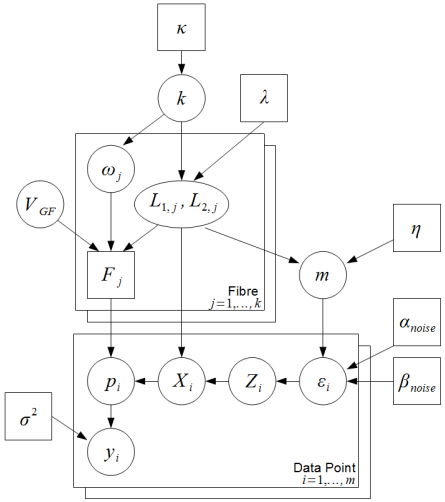


Warwick Statistics

# Our statistical model (IV)

## DAG of full model

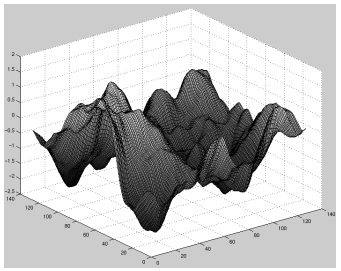







# Orientation Field

- Calculating an appropriate orientation field is key.
- Possible approaches include using random field theory, *eg* extending a Gaussian field . . .



- . . . but the configuration space of orientation fields is huge.
- Use Empirical Bayes to evade resulting problems.

# Empirical Bayes

- Allow prior for orientation field to depend on data.
- Directional evidence in data contributes to orientation field.

Empirical Bayes Step:

1. Partition data into noise and signal components
2. Calculate orientation fields for different partitions
   - Estimate local orientation at each point by nonlinear weighting (as used in Diffusion Tensor Imaging);
   - Interpolate orientation estimates using tensor means.
3. Various orientation fields acquire corresponding probabilities from partitions.

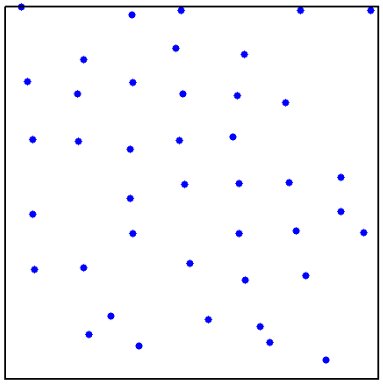
Warwick Statistics

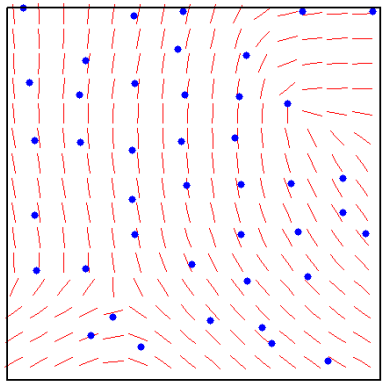




# Estimation of local orientation

- At point $x$, compute inertial tensor of transformations of the $y - x$ vectors using $\frac{y-x}{|y-x|} \exp\left(-\frac{1}{2\sigma^2}|y-x|^2\right)$.
- Use orientation of eigenvector for largest eigenvalue.
- Interpolate using weighted log-Euclidean tensor means.



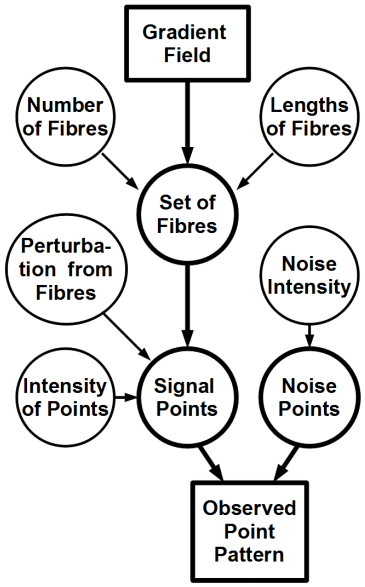


Warwick Statistics

# Sampling from Posterior Distribution of Fibres

Use continuous time Birth-Death Monte Carlo methods:

- Fibres are born and die
- Points allocated to clusters (fibres)
- Other moves:
  - Move fibres
  - Adjust lengths of fibres
  - Update individual parameters
  - Split and join fibres



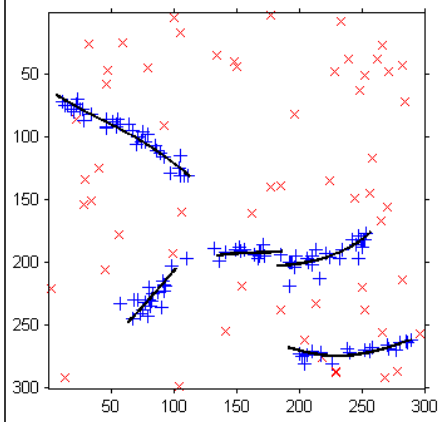

Warwick Statistics

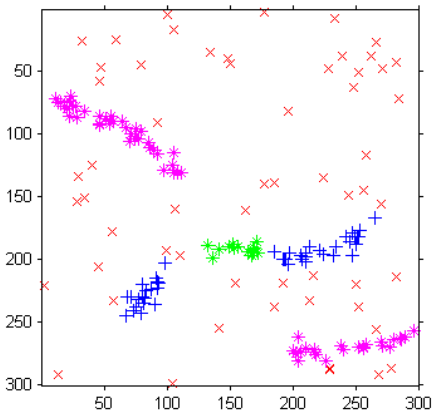




# Simulated Data (I)

Estimate of the clustering of the signal points

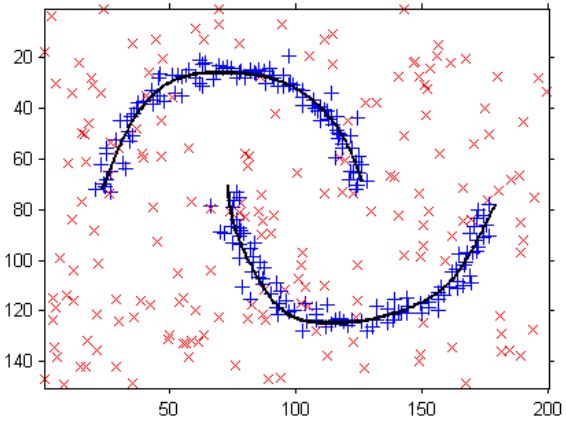


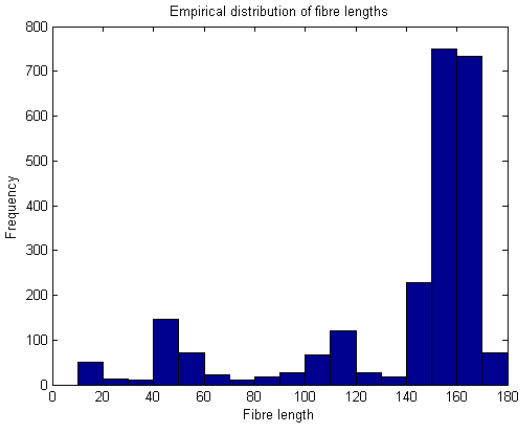




Warwick Statistics

# Simulated Data (II)

Example of a sample from the output



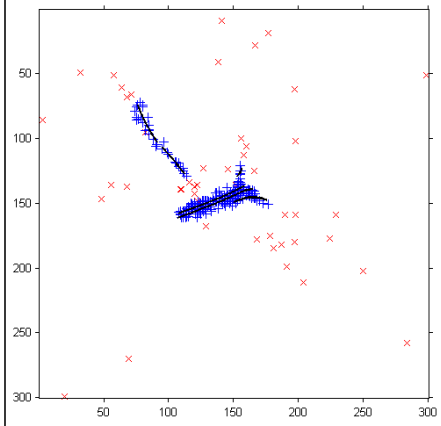



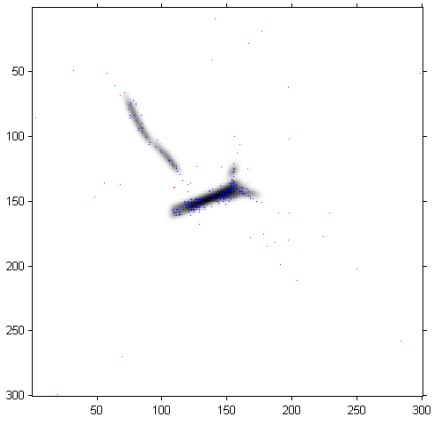





# Earthquake Data

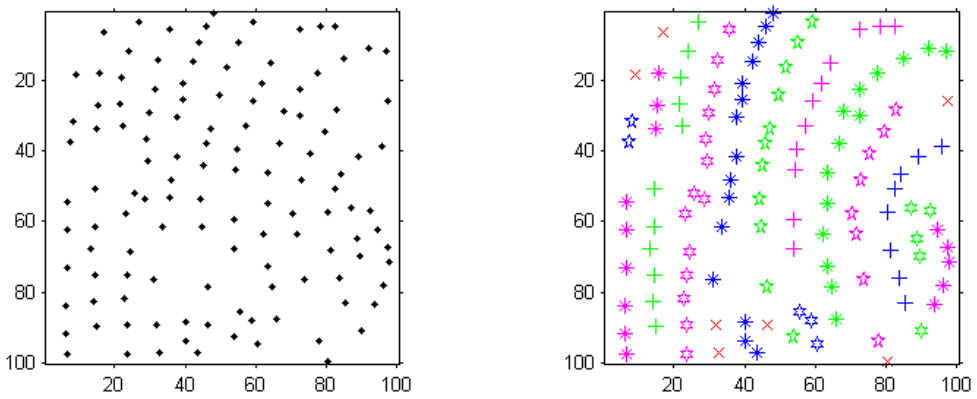
Estimate of the density of signal points



# Fingerprints

Estimate of clustering of signal points







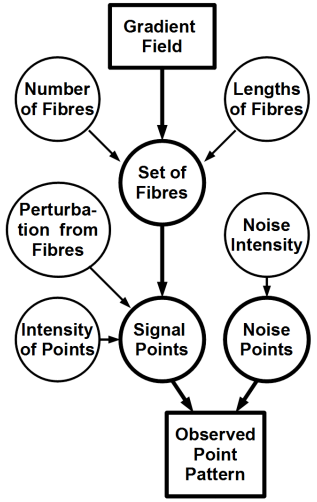
Warwick Statistics





# Conclusion

- Can identify fibre processes in point patterns with few constraints on fibre properties:
  - Identify fibres as streamlines of a gradient field;
  - Sample from posterior fibre distribution using BDMCMC;
  - Use Empirical Bayes prior for gradient field.
- Further Work:
  - Allow parameters to vary with fibres;
  - Address 3-dimensional problem (*eg:* galaxies).




Warwick Statistics

Arsigny, V., P. Fillard, X. Pennec, and N. Ayache (2006).
Log-Euclidean metrics for fast and simple calculus on diffusion tensors.
*Magnetic resonance in medicine 56*(2), 411–421.

Genovese, C. R., M. Perone-Pacifico, I. Verdinelli, and L. Wasserman (2009).
On the path density of a gradient field.
*The Annals of Statistics 37*(6A), 3236–3271.

Stanford, D. and A. Raftery (2000).
Finding curvilinear features in spatial point patterns: principal curve clustering with noise.
*IEEE Transactions on Pattern Analysis and Machine Intelligence 22*(6), 601–609.

Stoica, R. S., V. J. Martínez, and E. Saar (2007, August).
A three-dimensional object point process for detection of cosmic filaments.
*Journal of the Royal Statistical Society: Series C (Applied Statistics) 56*(4), 459–477.

Watson, C. (2001).
Dual Resolution Images from Paired Fingerprint Cards.