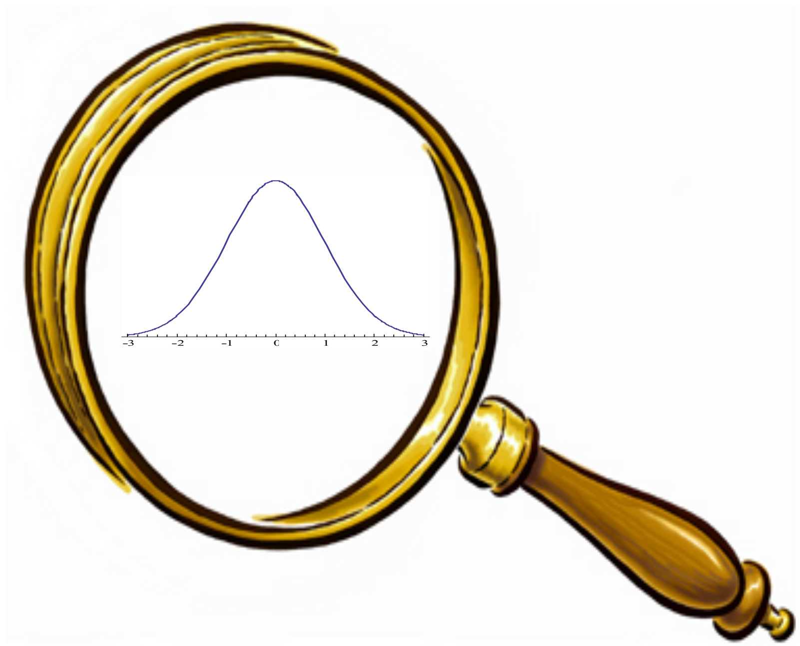
Quantitative Firearms and Toolmark Analysis:
New Developments and Software



Petraco Group
City University of New York, John Jay College

Outline

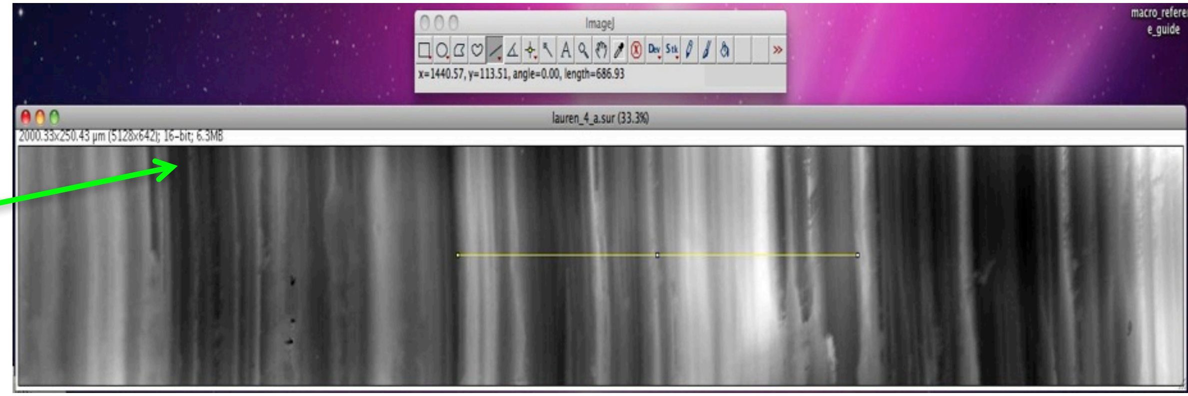
- 3D toolmark data, pre-processing and feature extraction: **x3pr**, **feature2**
- The statistics:
  - Identification Error Rates
  - “Match” confidence estimate from Conformal Prediction Theory: **cptID**
  - “Match” probability estimates from Empirical Bayes: **fdrID**
  - “Match” probability estimates from CMS data and Bayesian Networks

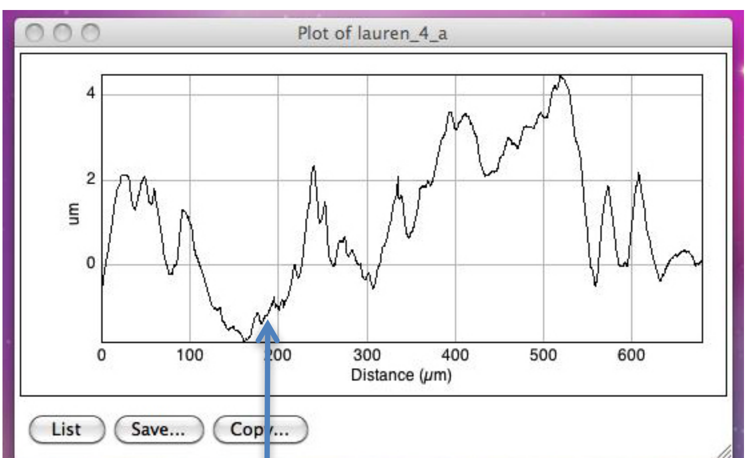
Data Acquisition For Toolmarks

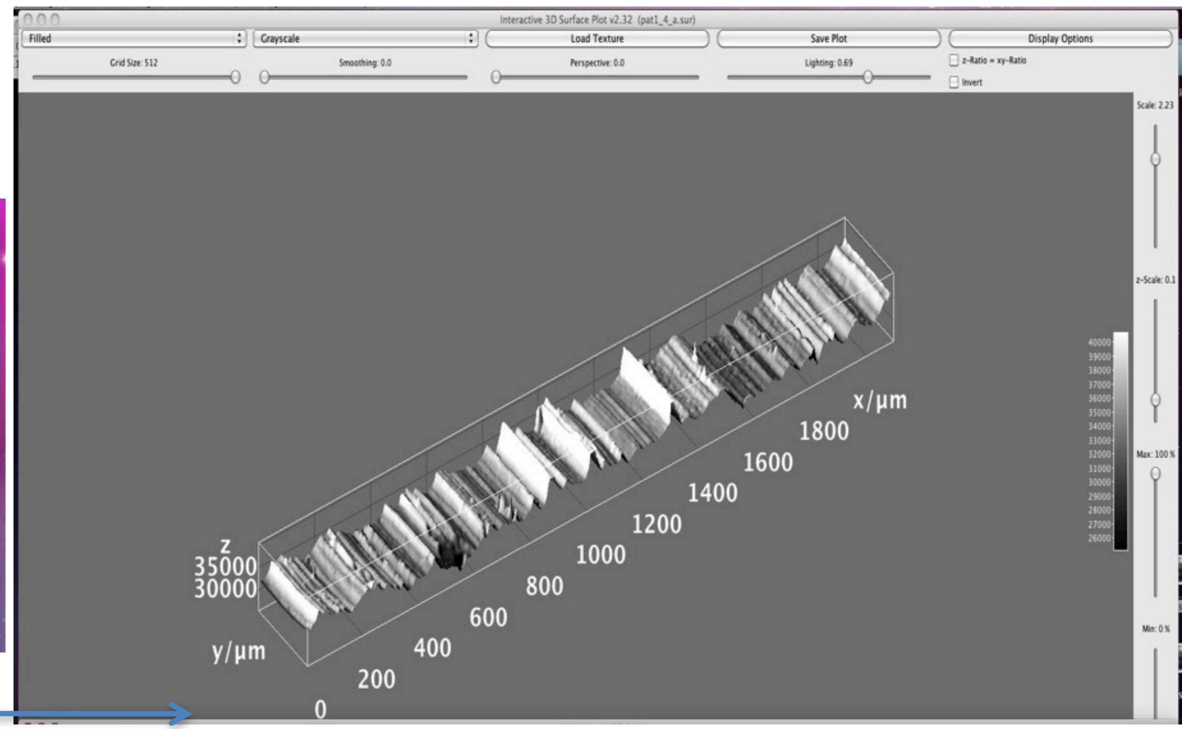
Confocal Microscope

Focus Variation Microscope

Screwdriver Striation Patterns in Lead



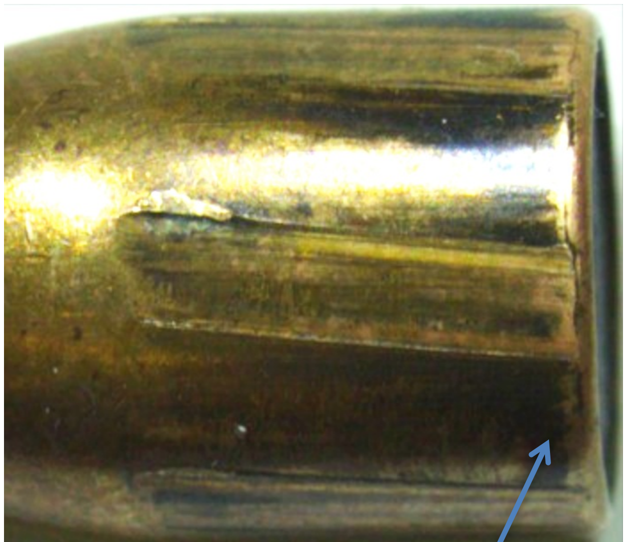




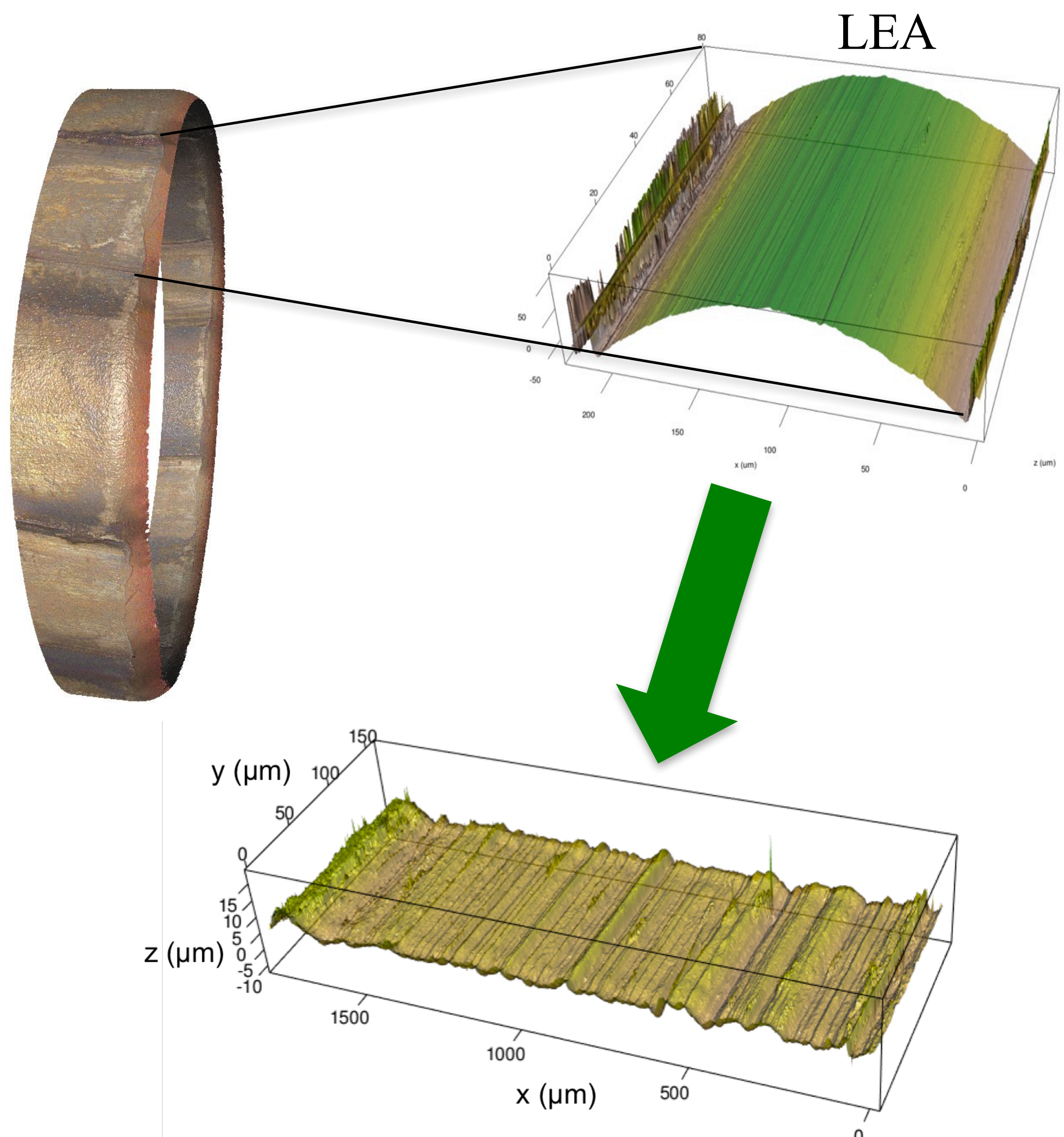
2D profiles

3D surfaces (interactive)

Bullets

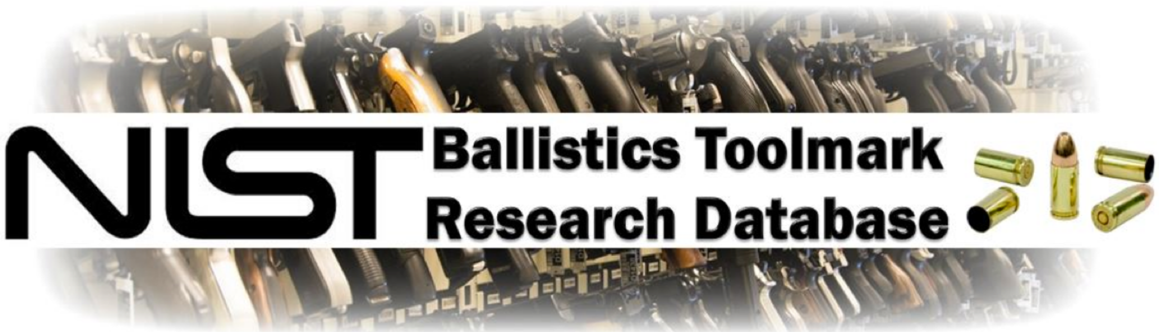


Bullet base, 9mm Ruger Barrel

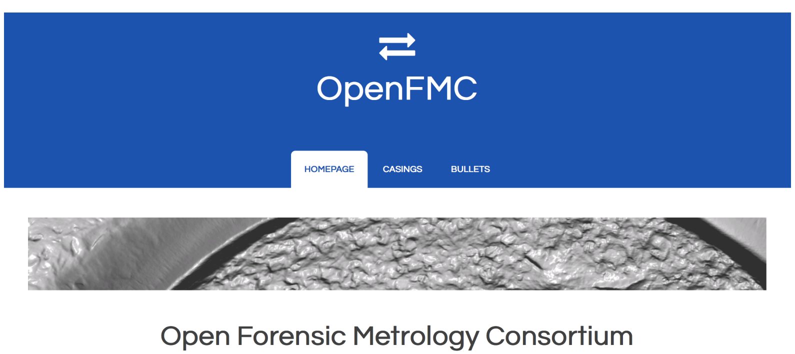


Toolmark Surface Data

- A growing database[Zheng]:
  - http://www.nist.gov/forensics/ballisticsdb/



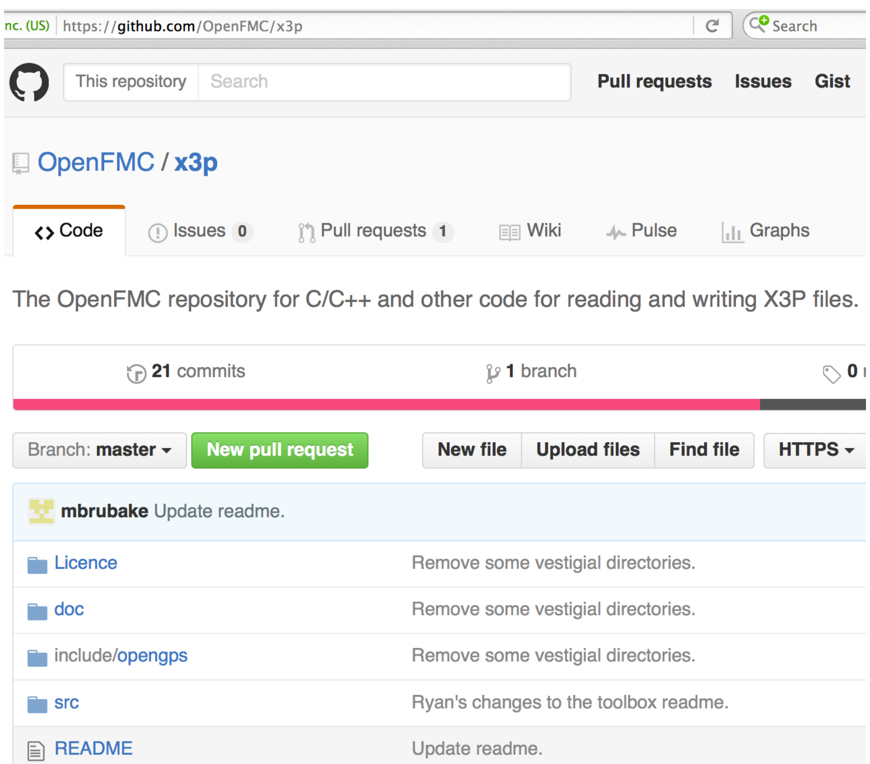
- Put in your two cents: OpenFMC[Lillien]
  - http://www.openfmc.org/

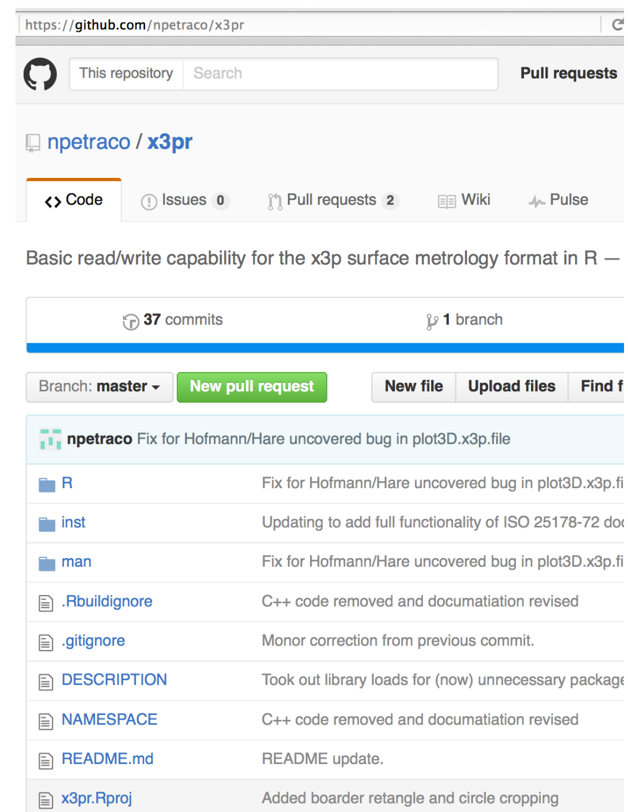


Toolmark Surface Data

- Standardizing file format: .x3p
  - http://www.nist.gov/forensics/ballisticsdb/dataformat.cfm

x3p C++/ library, Windows[Brubaker]   **x3pr**[Petraco] for R, Any OS

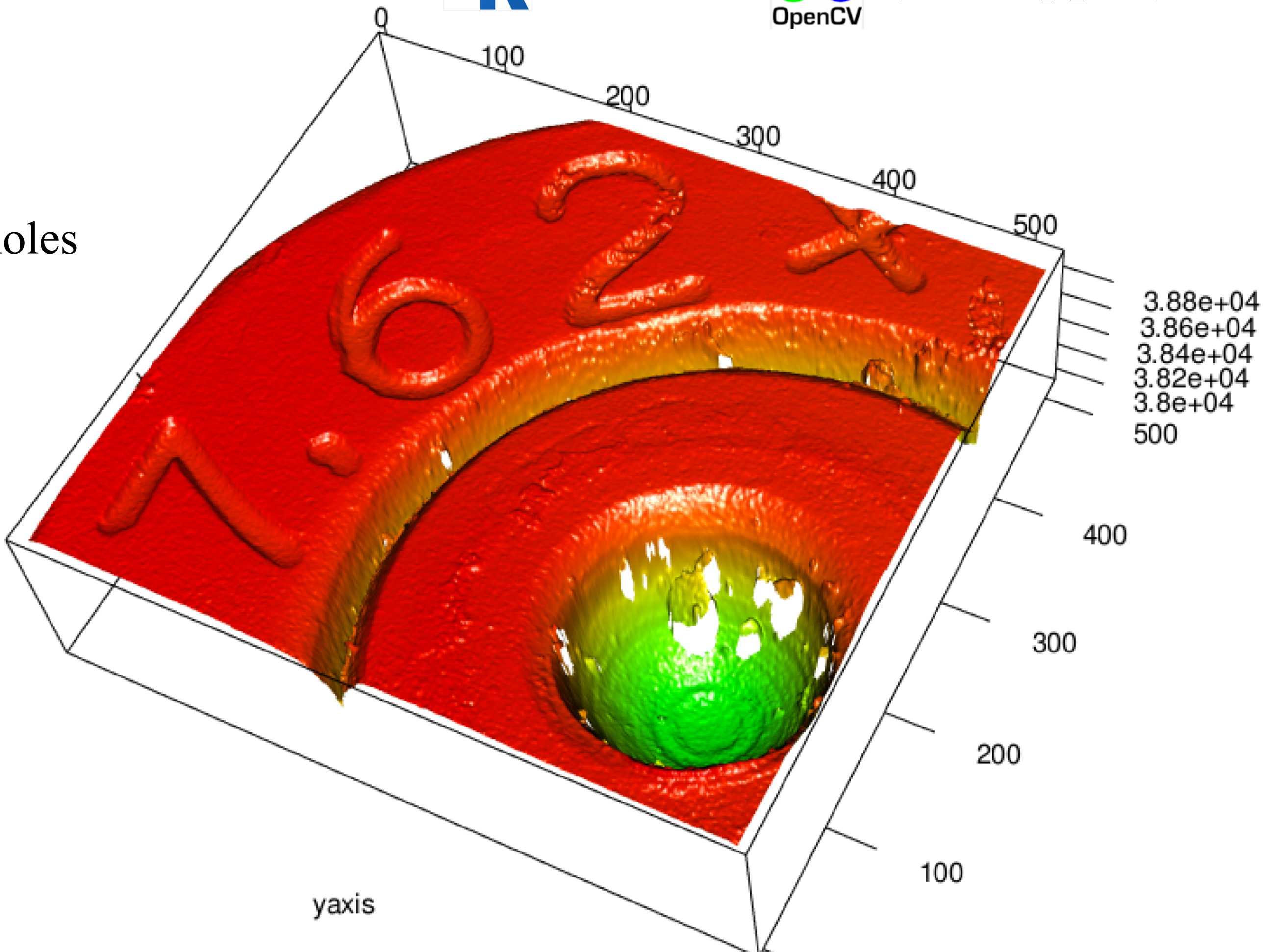




Pre-processing Surface Data

- 3D tool mark data usually needs (a lot of…) preprocessing
  - We use a combination of R and **C++**/ OpenCV (via **Rcpp**[Edd]) = **feature2**[Petraco]

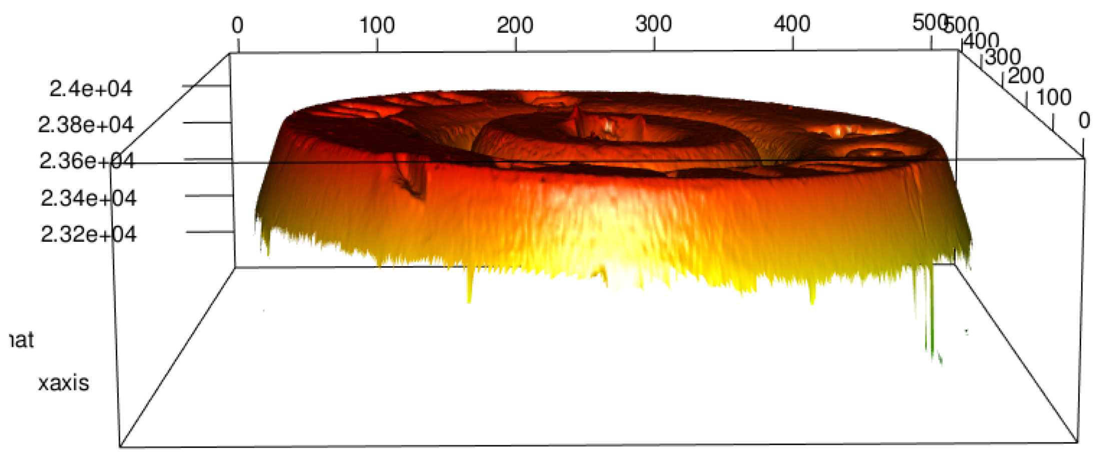
Possibly fill holes



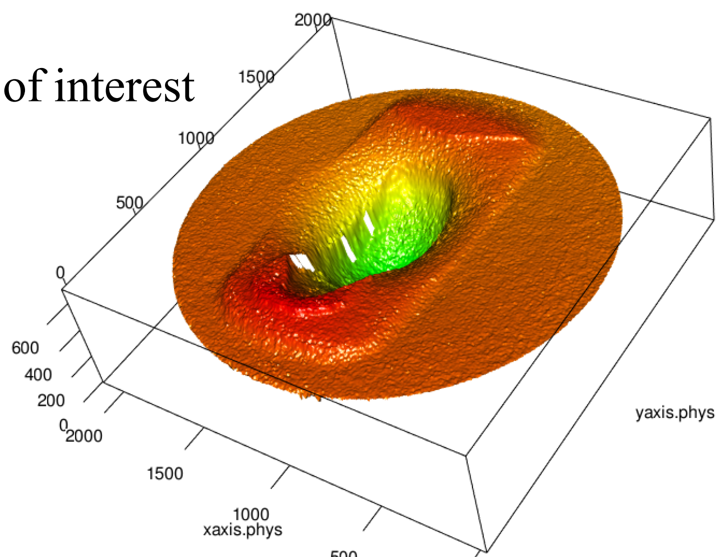
Preprocessing Surface Data

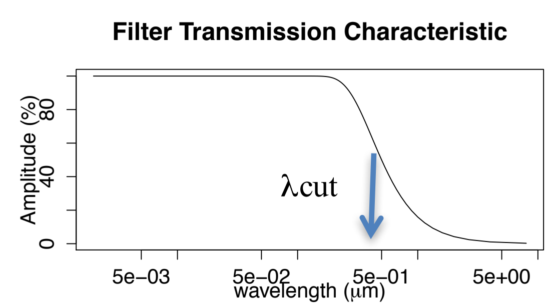
- 3D tool mark data usually needs (a lot of…) preprocessing
  - In **feature2**:

Possibly remove "long range" behavior (leveling, form removal)



Crop out areas of interest





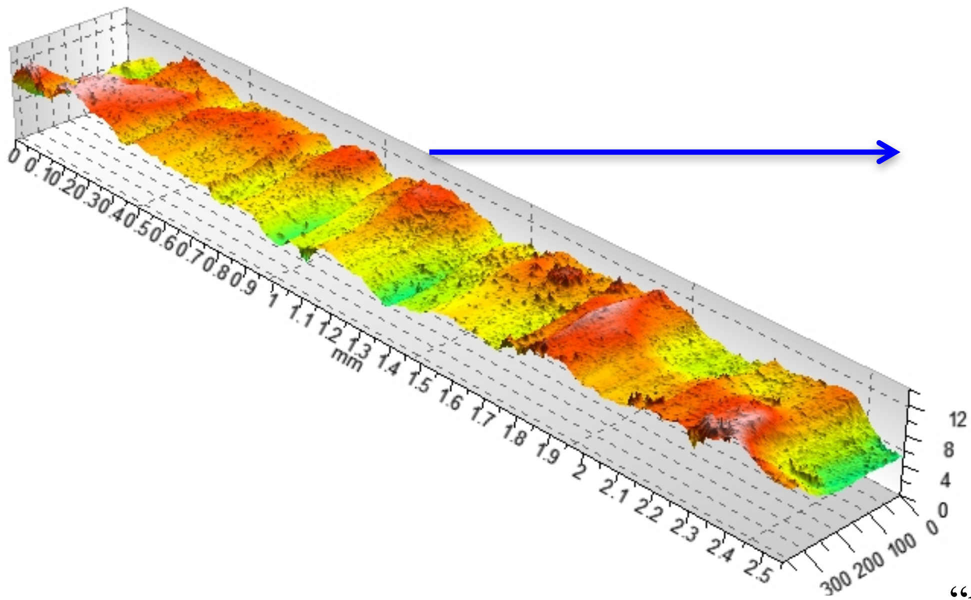
Bandpass filters via:

waveslim Whitcher

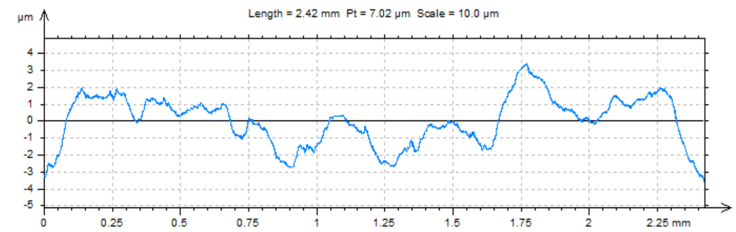
Good Features are the Key!

- We need a tool mark feature set that is:
  - Large in number
  - (possibly) transnationally invariant
  - (possibly) rotationally invariant
  - Mostly statistically independent
  - DISCRIMINATORY!

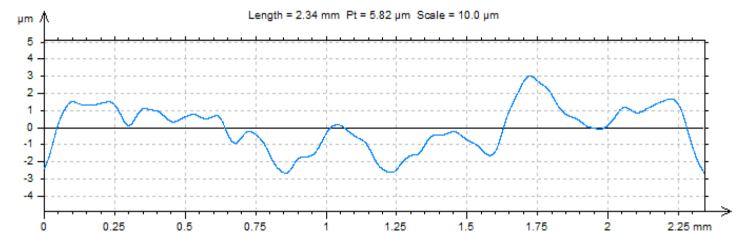
Toolmark Features



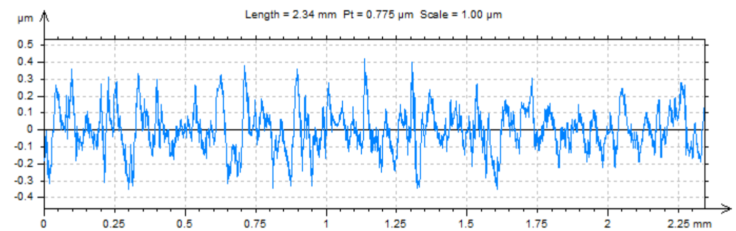
Mean total profile:



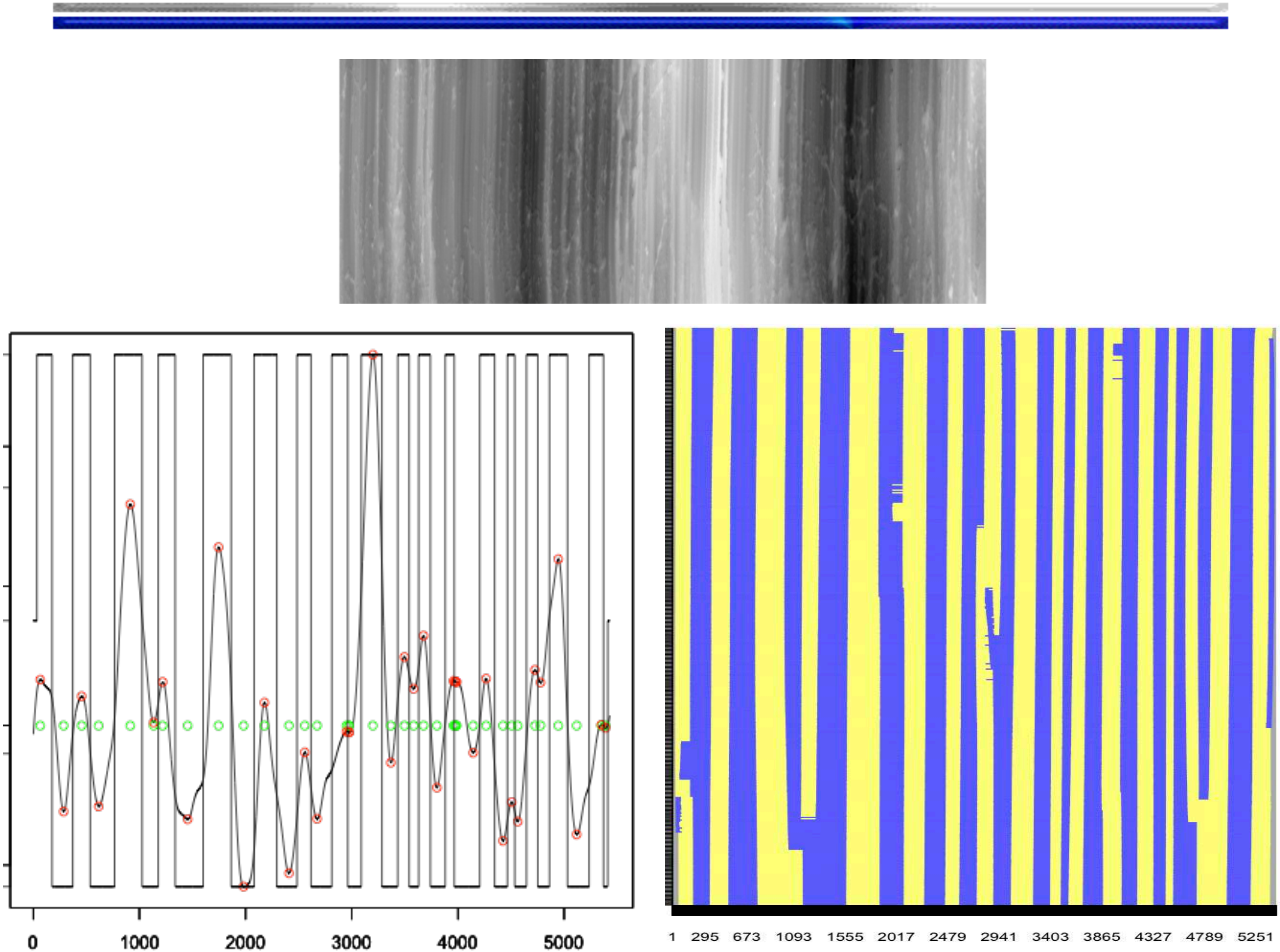
Mean "waviness" profile:



Mean "roughness" profile:



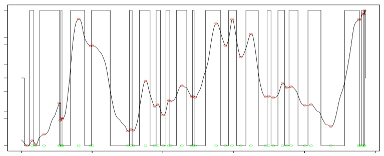
Aperture primer shear on a 9mm cartridge case fired from the a Glock 19


0 1000 2000 3000 4000 5000
1 295 673 1093 1555 2017 2479 2941 3403 3865 4327 4789 5251

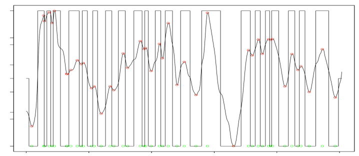
FAST-Consecutive Matching Striae (CMS)-Space

Biasotti-Murdock dictionary: “Closest Match Ref Set”

Rcpp and **parallel**[R-core] packages are great for quick and easy speedups

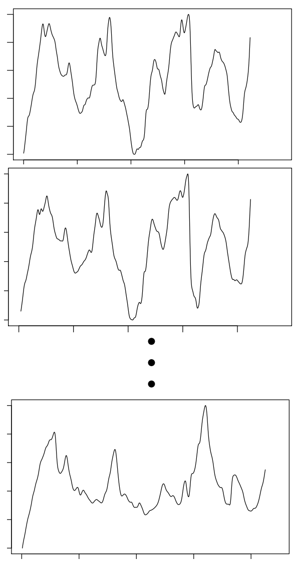


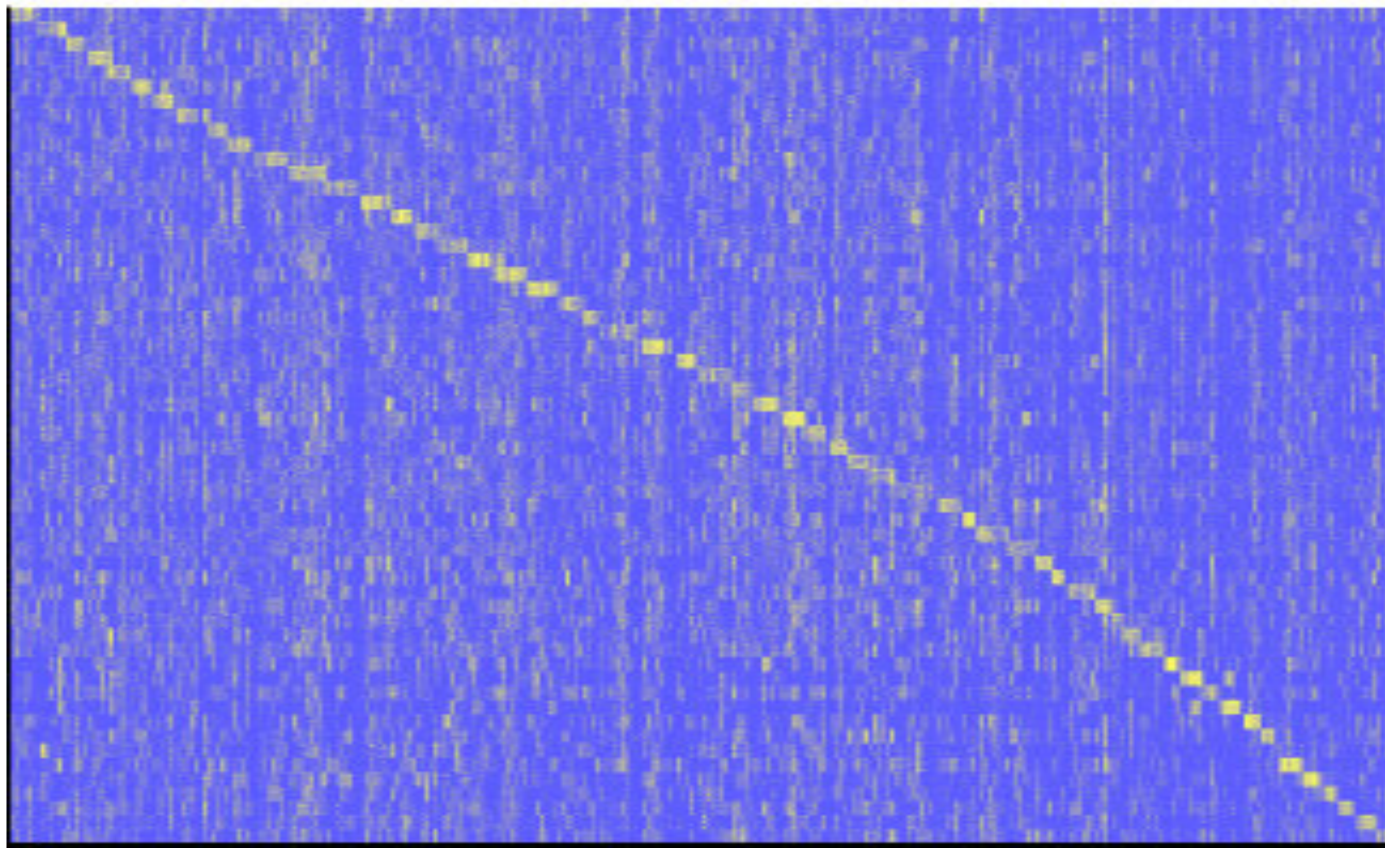
… … …



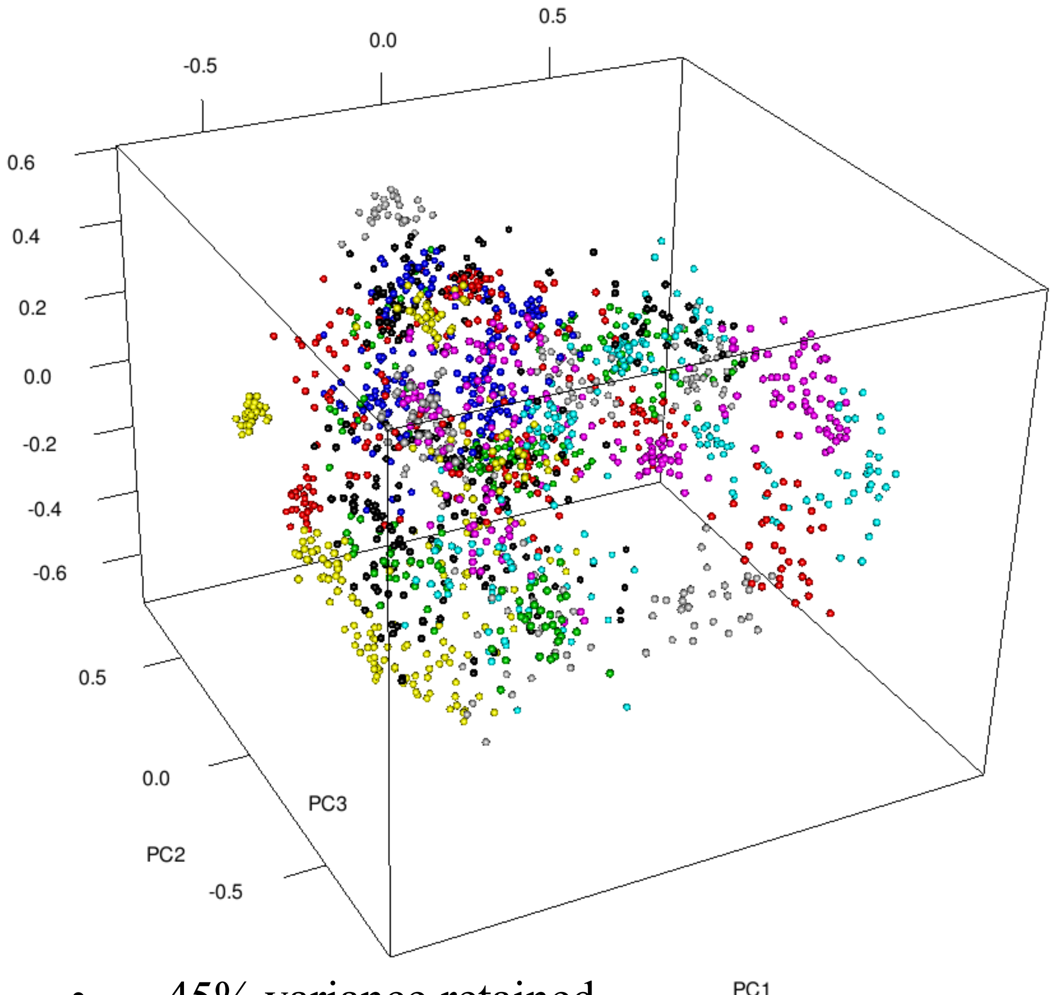
… … …

Database/queries





- 3D PCA of 1740 real and simulated mean profiles of striation patterns from 58 screwdrivers:



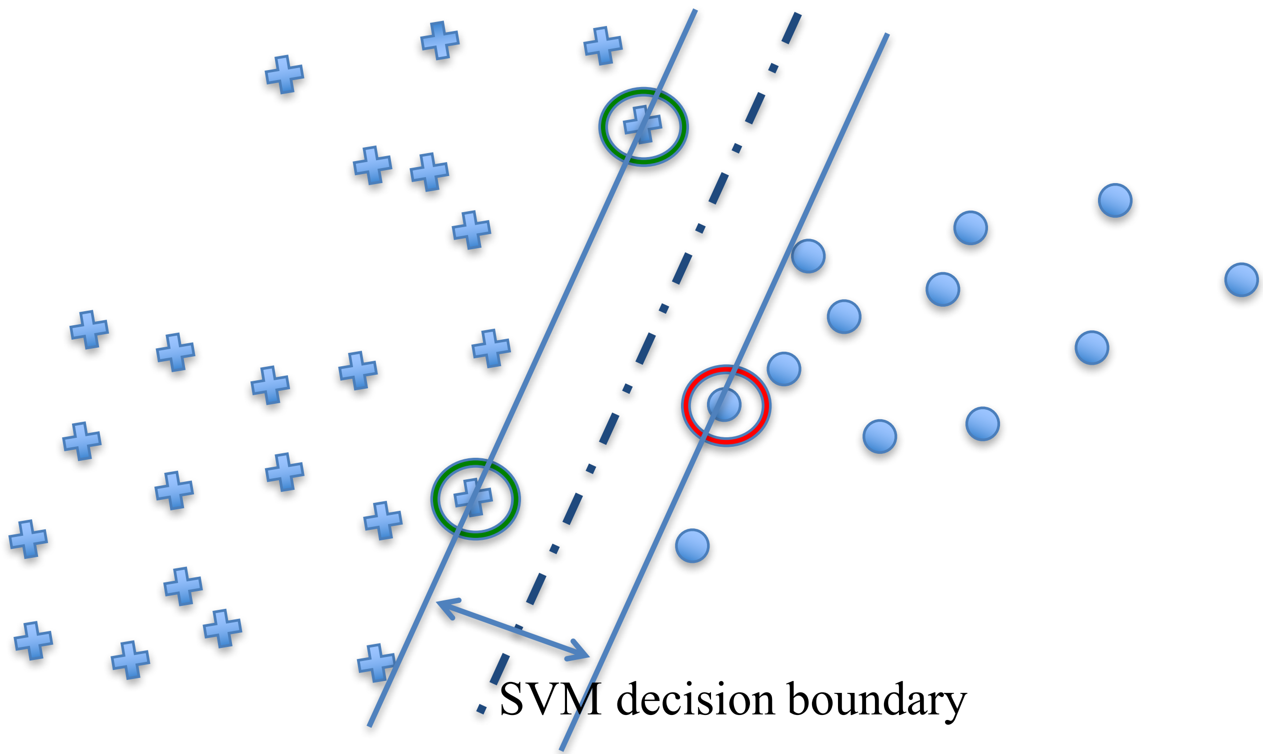
- How many PCs should we use to represent the data??
    - No unique answer

- FIRST we ***need an algorithm to I.D. a toolmark*** to a tool

- ~45% variance retained

Support Vector Machines

- Support Vector Machines (SVM) determine efficient association rules
  - ***<u>In the absence of specific knowledge of probability densities</u>***

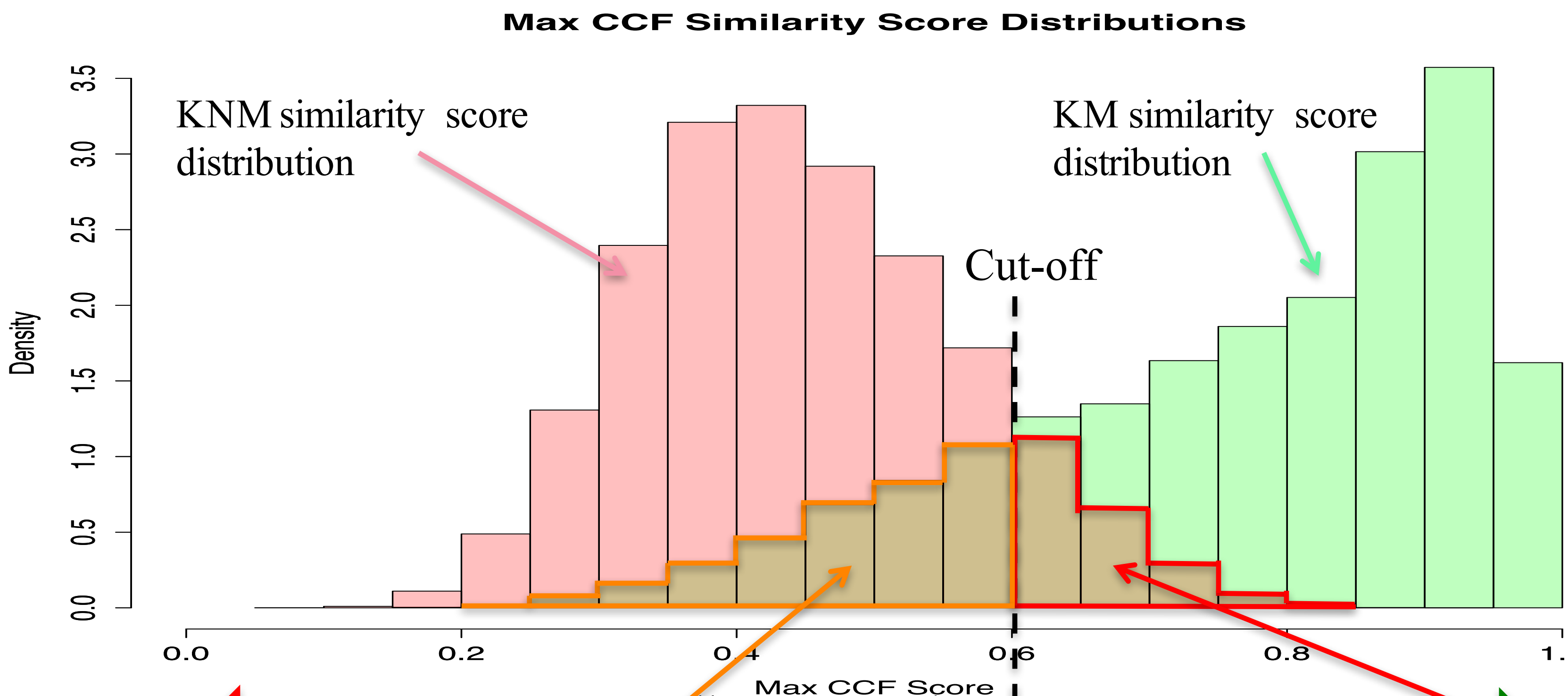


Error Rate Estimation: Machine Learning

- **Cross-Validation**: hold-out chunks of data set for testing
    - Known since 1940s
    - Most common: **Hold-one-out**
- **Bootstrap**: Randomly selection of observed data (with replacement)
    - Known since the 1970s
    - Can yield <u>***confidence intervals around error rate estimate***</u>
- **The Best**: Small training set, BIG test set

Error Rate Estimation: Pair-Wise Comparisons

- **Univariate** approaches compute estimates of similarity score distributions for **Known Matches** (KM) and **Known Non-Matches** (KNM)



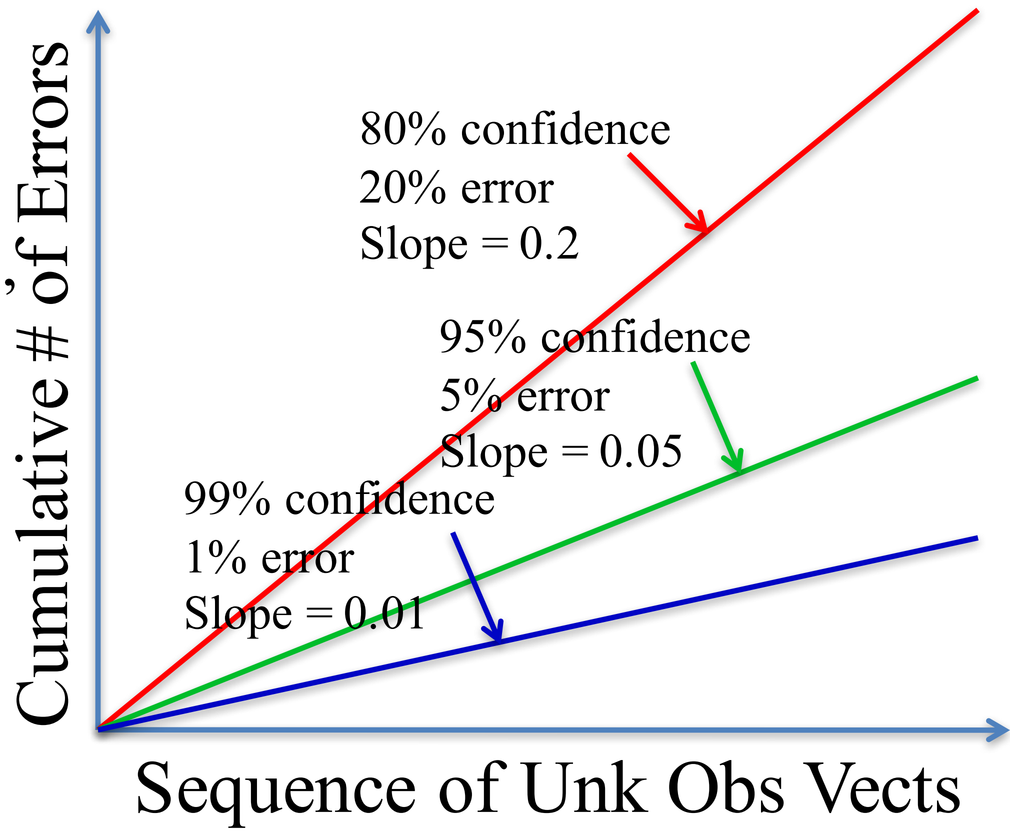
#False Non-matches/#KM-Comparisons = $\widehat{\text{FNMR}}$ #False Matches/#KNM-Comparisons = $\widehat{\text{FMR}}$

$\widehat{\text{Error Rate}}$ = (#False Non-matches + #False Matches)/#Comparisons

How good of a "match" is it?

Conformal Prediction^Vovk

- Can give a judge or jury an easy to understand measure of reliability of classification result
  - *Confidence* on a scale of 0%-100%
  - *Testable claim*: Long run I.D. error-rate should be the chosen significance level
- This is an orthodox "frequentist" approach
  - Roots in Algorithmic Information Theory
- Data should be IID but that's it



How Conformal Prediction works for us

- Given a “bag” of obs with known identities and one obs of unknown identity[Vovk]
  - Estimate how “wrong” labelings are for each observation with a **non-conformity** score (“wrong-iness”)
  - For us, one-vs-one SVMs: $t_i = \frac{1}{k-1} \sum_{j=1}^{k(k-1)/2} \lambda_{i,j}$

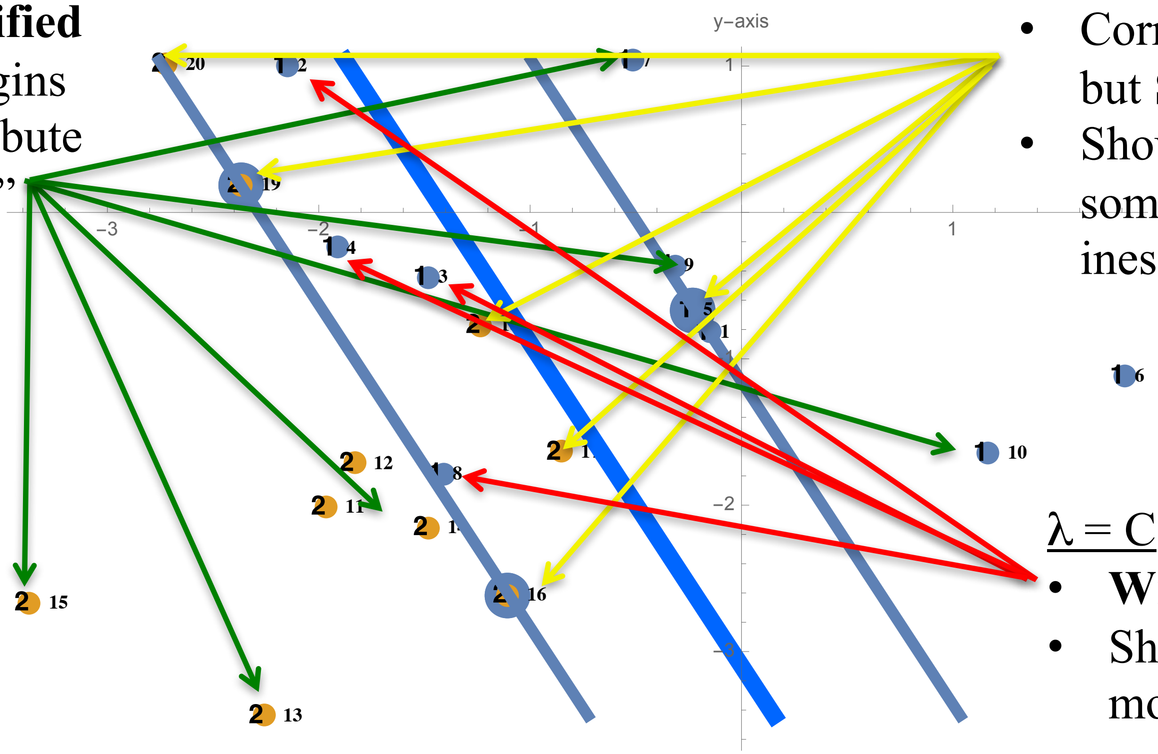
<u>$\lambda = 0$</u>

- **Correctly classified** and behind margins
- Shouldn’t contribute to “wrong-iness”

<u>$0 < \lambda < C$</u>

- Correctly classified but **SVs or marginal**
- Should contribute something to “wrong-iness”

Intuition:



<u>$\lambda = C$</u>

- **Wrong**
- Should contribute most “wrong-iness”

How Conformal Prediction works for us

- Given a “bag” of obs with known identities and one obs of unknown identity[Vovk]
  - Looking at the “wrong-iness” for all the known observations in the bag:
    - Ask: Does labeling-*i* for the unknown have an unusual amount of “wrong-iness”??:

Given “wrong-iness” for labeling-*i* of unknown, number of obs with at least as much “wrongi-ness”

$$p_{\text{possible-ID}_i} = \frac{\#\left\{ j \in \{1,2,\ldots,n\} : t_j^{\text{possible-ID}_i} \geq t_{\text{test-pattern}}^{\text{possible-ID}_i} \right\}}{n} \qquad i \in \{1,2,\ldots,k \text{ I.D.s}\}$$

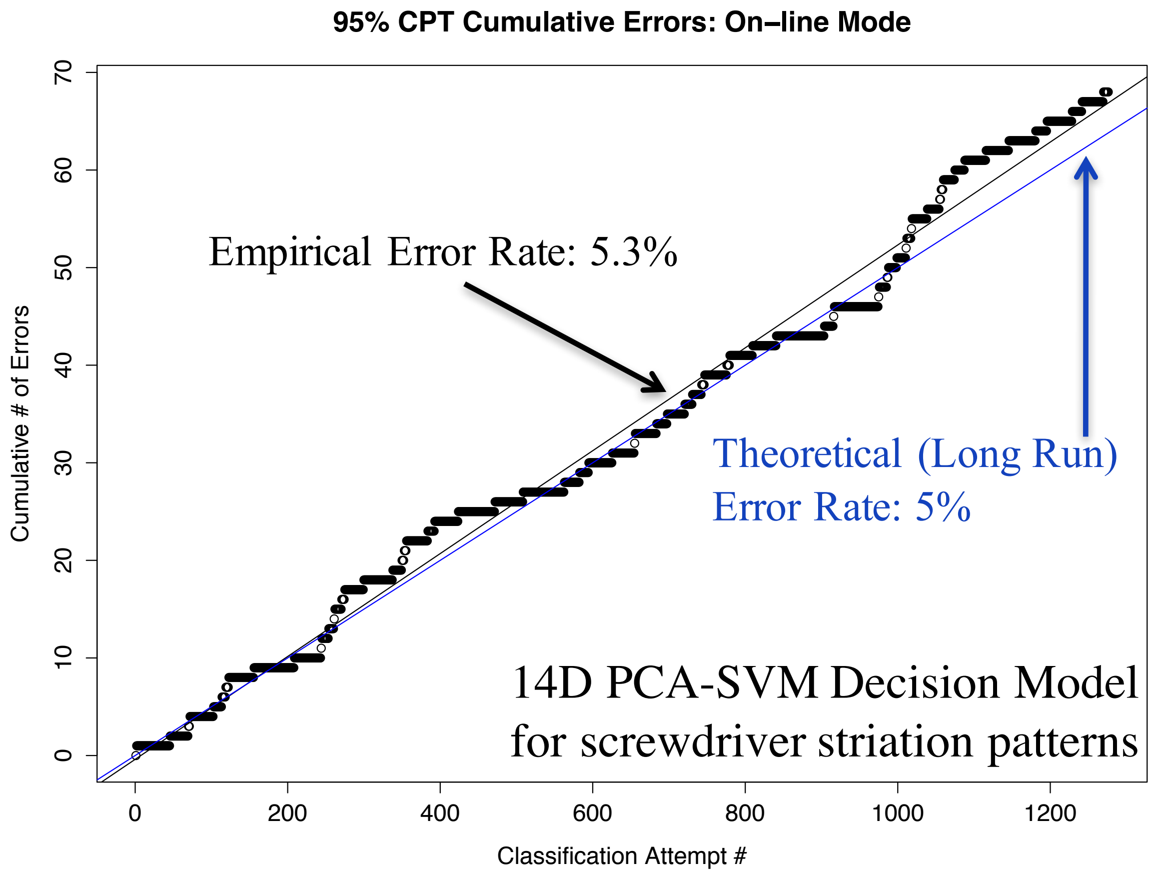
- If not:
  - $p_{\text{possible-ID}_i} \geq$ chosen level of significance α
  - Put ID_i in the $(1 - \alpha)$*100% confidence interval: $\Gamma^{1-\alpha}$

$$\text{ID}_i \in \Gamma^{1-\alpha} \quad \text{if } p_{\text{ID}_i} \geq \alpha$$

Conformal Prediction

- For 95%-CPT (PCA-SVM) confidence intervals will not contain the correct I.D. 5% of the time *in the long run*
  - *Straight-forward validation/explanation picture for court*

cptIDPetraco for R
Coming soon…



How good of a “match” is it?
Efron Empirical Bayes

- An I.D. is output for each questioned tool mark
  - This is a computer “match”
- What’s the **probability the tool is truly the source of the tool mark**?
- Similar problem in genomics for detecting disease from microarray data
  - They use *data* **and** *Bayes’ theorem* to get an estimate

Empirical Bayes

- From Bayes' Theorem we can get[Efron]:

Estimated probability of not a true "match" given the algorithms' output z-score associated with its "match" →

$$\hat{\Pr}\left(S^{-} \mid z\right) = \frac{\hat{p}\left(z \mid S^{-}\right)}{\hat{f}\left(z\right)} \hat{\Pr}\left(S^{-}\right)$$

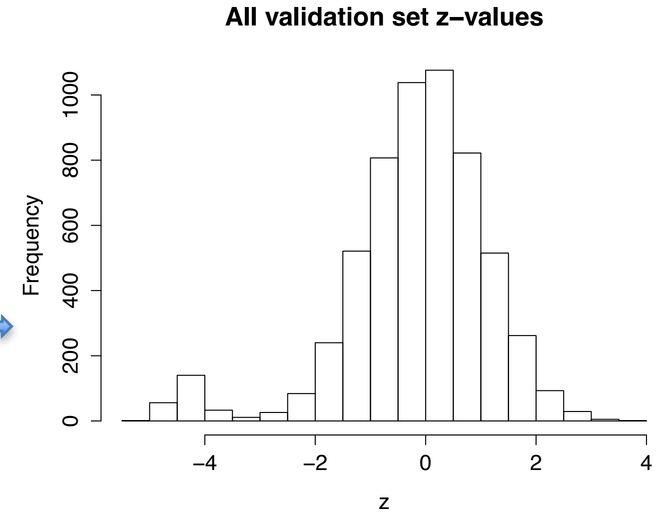
Names: **Posterior error probability (PEP)**[Kall]
Local false discovery rate (lfdr)[Efron]

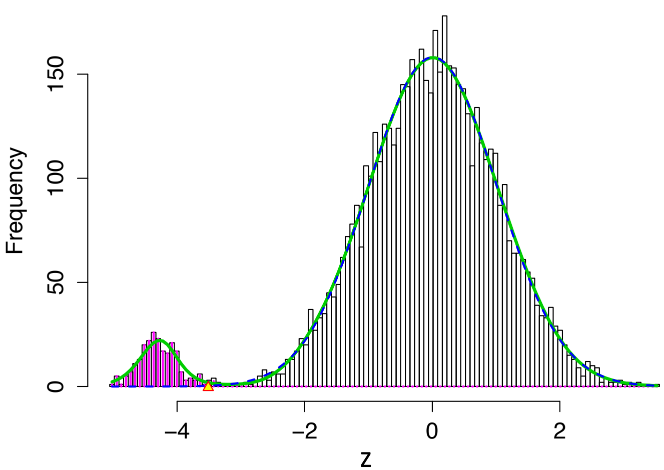
- Suggested interpretation for casework:

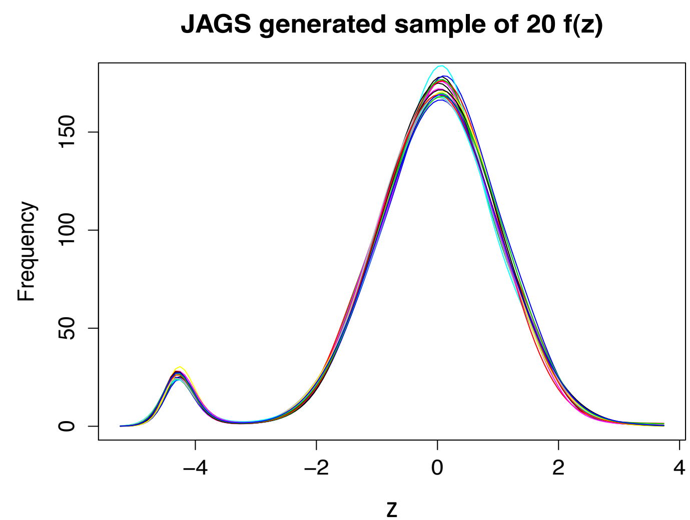
$1 - \hat{\Pr}\left(S^{-} \mid z\right)$ = Estimated "***believability***" that the specific tool produced the tool mark

Fit local-fdr models

$\Phi^{-1}(p\text{-values})$







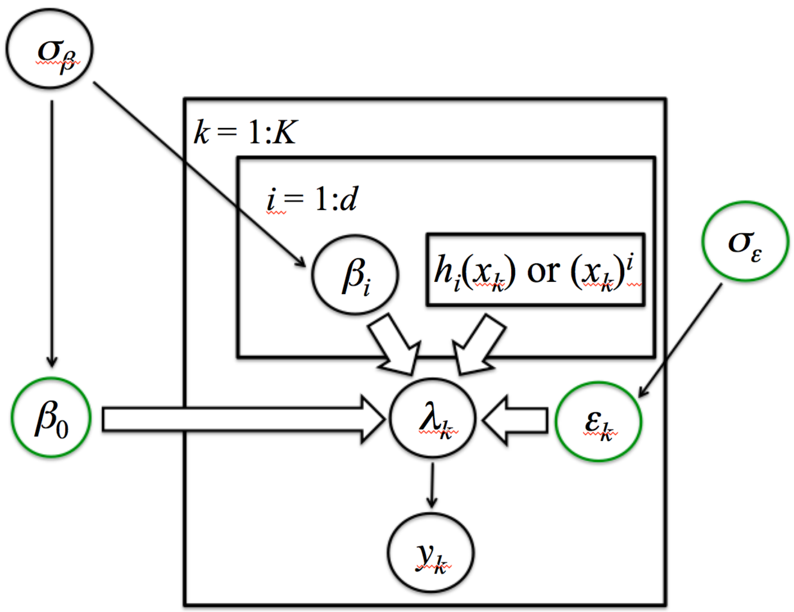
Use locfdr[locfdr]
Fit classic Poisson regression for f(z)

Use modified locfdr/JAGS[JAGS,Plummer] or Stan[Stan]
Fit Bayesian hierarchical Poisson regressions

Bayesian Hierarchical Poisson Regression Details

- To run the Bayesian Estimation we use **JAGS**[Plummer] or **Stan**[Gelman]:

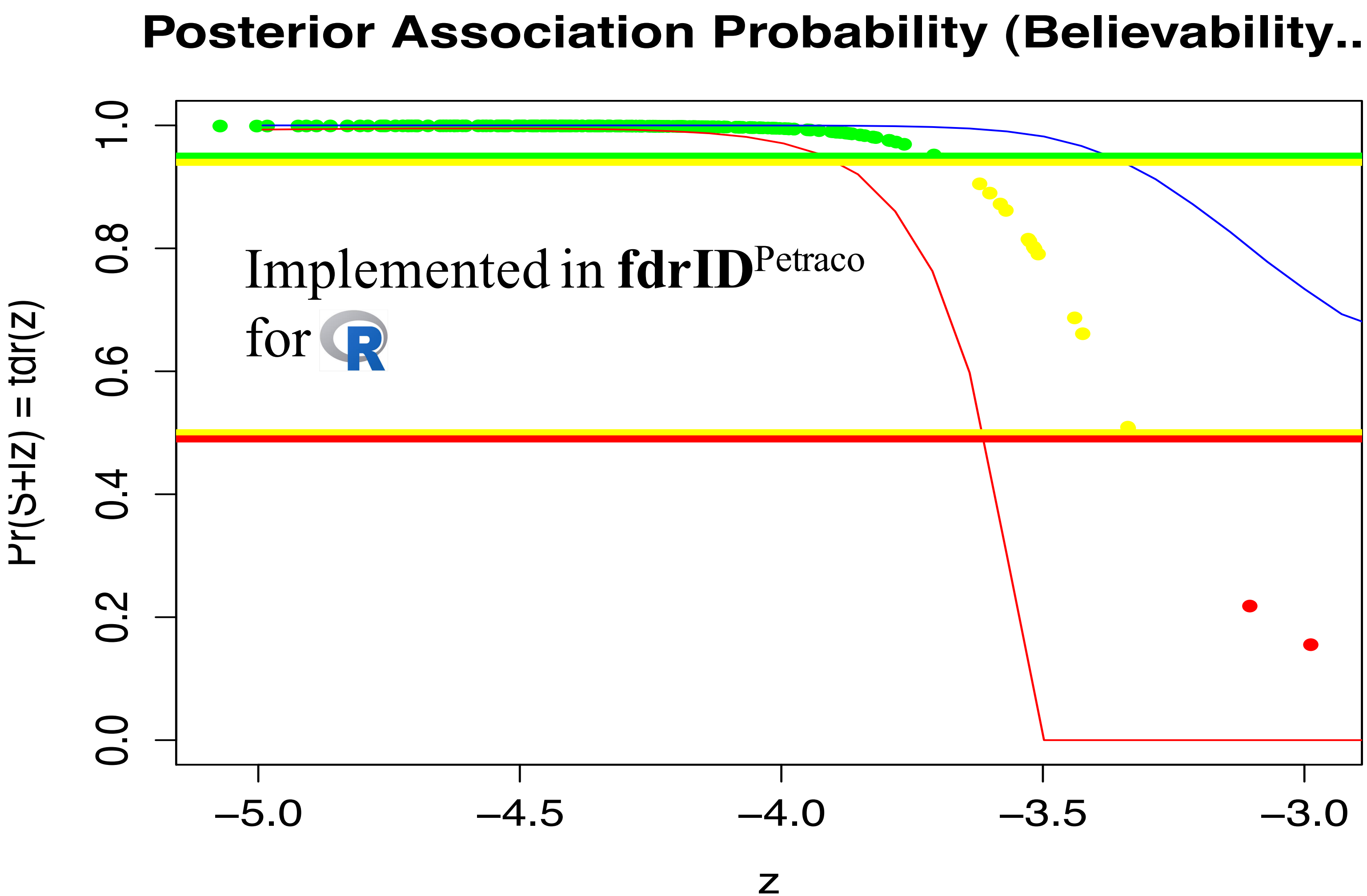
DAG for the Poisson Regression



$$\begin{aligned}
\sigma_\beta &\sim \mathrm{Uniform}(0, 100) \\
\sigma_\epsilon &\sim \mathrm{Uniform}(0, 100) \\
\beta_0 &\sim \mathrm{Normal}(0, \sigma_\beta) \\
\epsilon_i &\sim \mathrm{Normal}(0, \sigma_\epsilon) \\
\beta_j &\sim \mathrm{Normal}(0, \sigma_\beta)
\end{aligned}$$

Suggested by Gelman

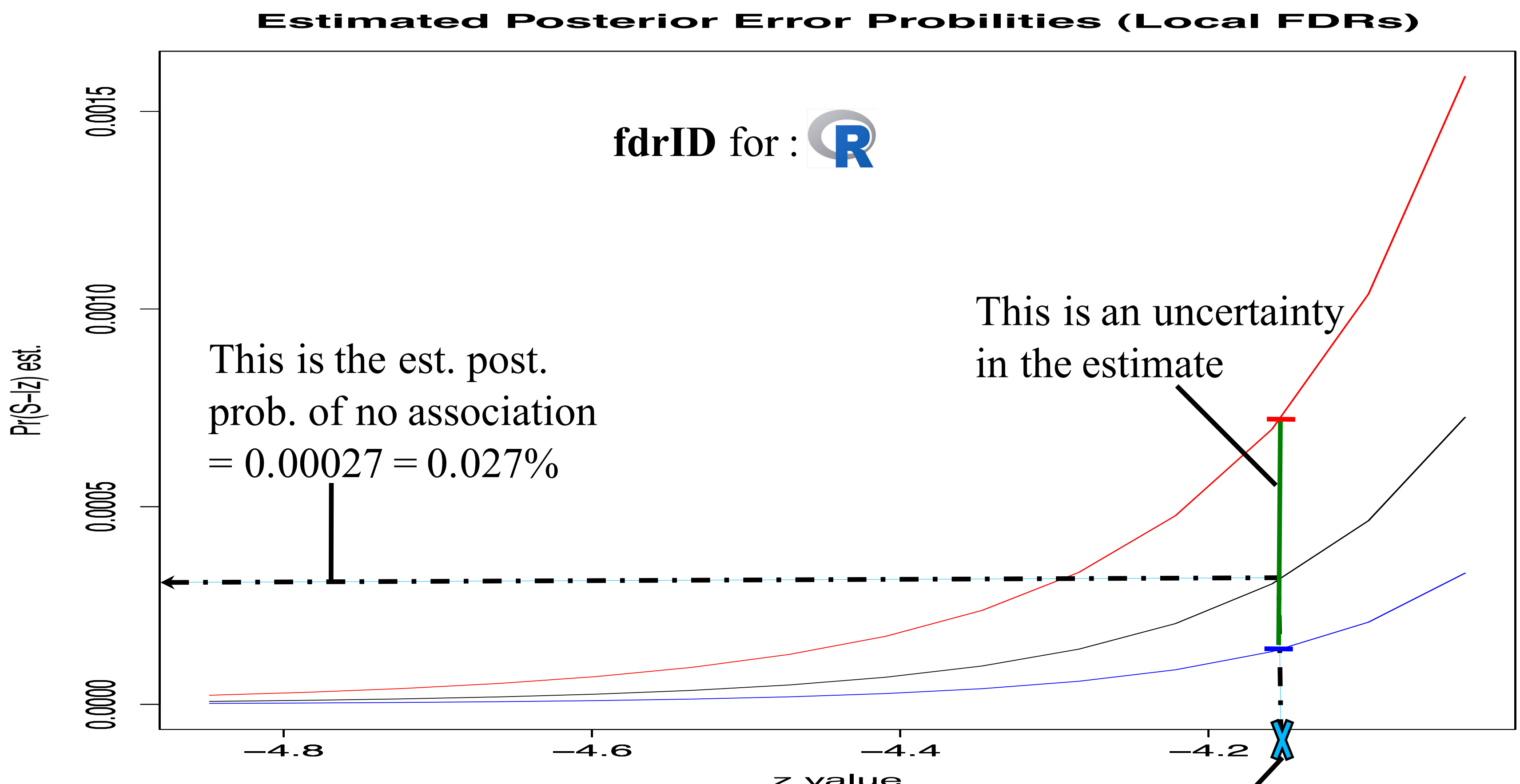
A Bayesian Hierarchical Model: Believability Curve



JAGS MCMC Bayesian over-dispersed Poisson with intercept, on test set

Empirical Bayes'

- Model's use with crime scene "unknowns":



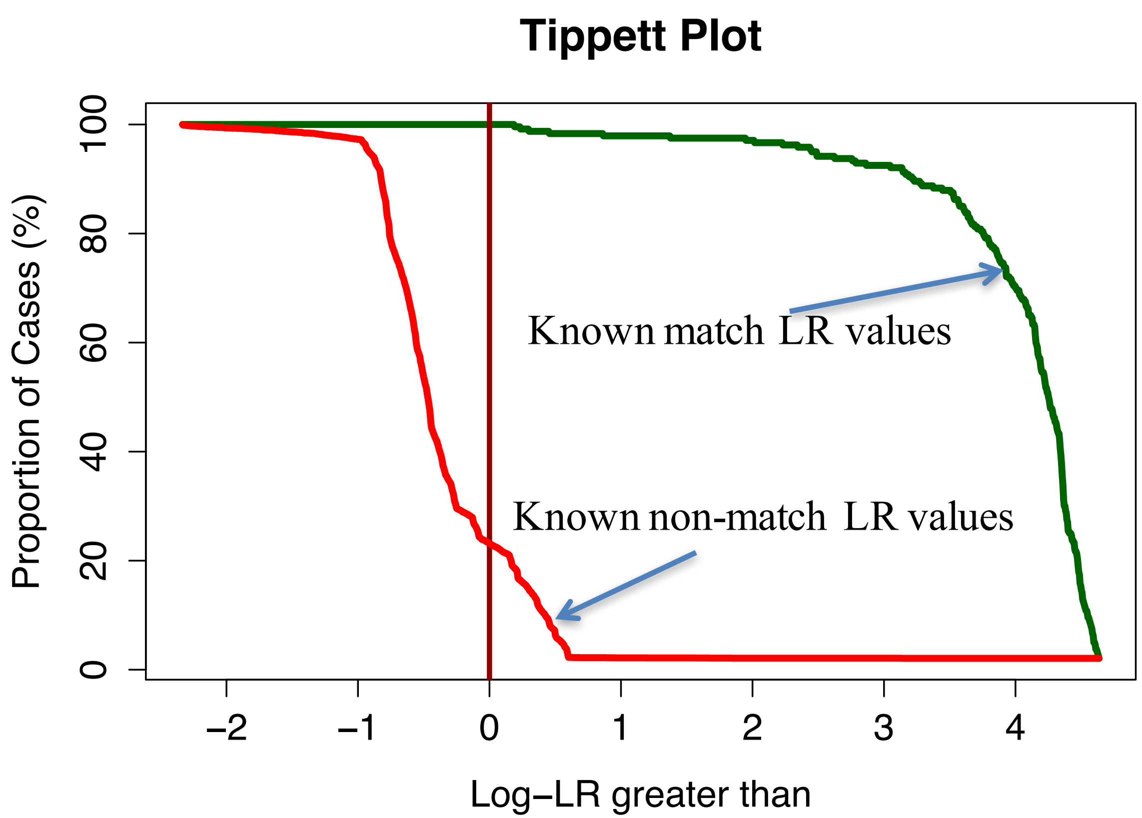
Computer outputs "match" for:

unknown crime scene toolmarks-with knowns from "Bob the burglar" tools

Likelihood Ratios from Empirical Bayes

- Using the fit posteriors and priors we can obtain the likelihood ratios

$$\widehat{\mathrm{LR}}(z) = \frac{\hat{\Pr}(H_p|E)}{\hat{\Pr}(H_d|E)} \Bigg/ \frac{\hat{\Pr}(H_p)}{\hat{\Pr}(H_d)} = \frac{\widehat{\mathrm{tdr}}(z)\ \hat{\pi}_0}{\widehat{\mathrm{fdr}}(z)\ 1 - \hat{\pi}_0}$$



Bayesian Match Probabilities from CMS

- 2007 Neel and Wells study[Neel, Wevers, Buckleton]:
    - Count the number of each type of CMS run for KM and KNM comparisons
        - A CMS type is its run length:
            - 4X means 4 matching adjacent lines in a comparison of two striation patterns

914 KM comparisons

Number observed	CMS run lengths: 2X	3X	4X	…
0	508	612	694	…
1	186	172	135	…
2	109	59	43	…
3	39	29	19	…
4	21	15	16	…
5	10	9	2	…
6	4	9	1	…
7	10	6	3	…
8	14	2	0	…
>8	13	1	1	…

1411 KNM comparisons

Number observed	CMS run lengths: 2X	3X	4X	…
0	771	1239	1357	…
1	298	124	47	…
2	143	35	4	…
3	84	10	2	…
4	46	2	1	…
5	21	1	0	…
6	13	0	0	…
7	14	0	0	…
8	6	0	0	…
>8	15	0	0	…

Model each column of counts as arising from a multinomial distribution with Dirichlet prior

Bayesian Match Probabilities from CMS

- Updated CMS run length probabilities:

KM comparisons

Number observed	CMS run lengths: 2X	3X	4X	…
0	0.550	0.663	0.752	…
1	0.202	0.187	0.147	…
2	0.119	0.065	0.047	…
3	0.043	0.032	0.022	…
4	0.024	0.018	0.019	…
5	0.012	0.011	0.003	…
6	0.005	0.011	0.002	…
7	0.012	0.008	0.004	…
8	0.016	0.003	0.001	…
>8	0.015	0.002	0.002	…

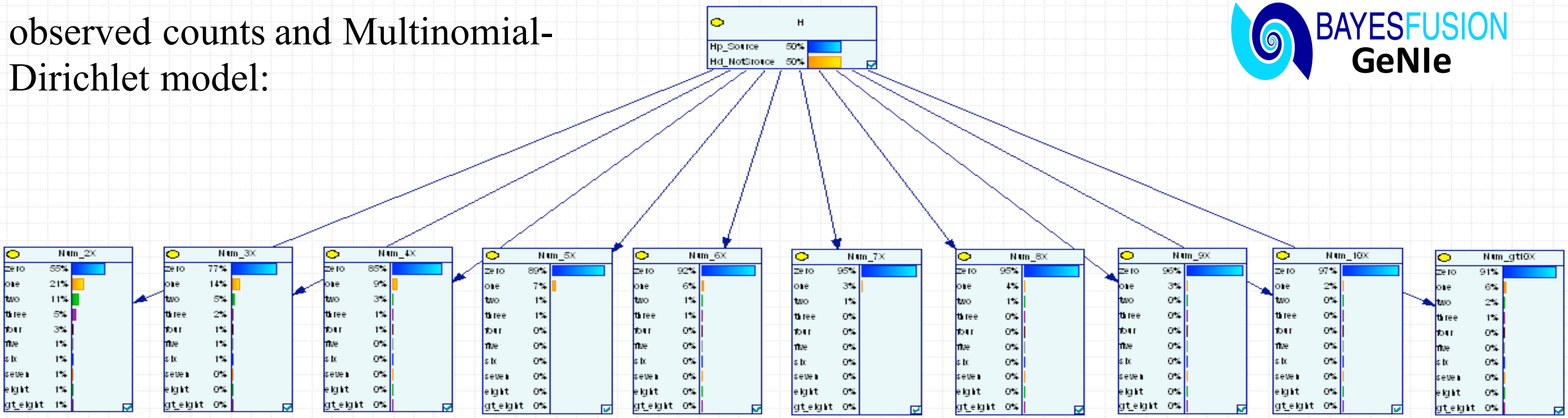
KNM comparisons

Number observed	CMS run lengths: 2X	3X	4X	…
0	0.5440	0.8726	0.9556	…
1	0.2099	0.0880	0.0338	…
2	0.1010	0.0254	0.0035	…
3	0.0598	0.0078	0.0021	…
4	0.0332	0.0021	0.0014	…
5	0.0155	0.0014	0.0007	…
6	0.0099	0.0007	0.0007	…
7	0.0105	0.0007	0.0007	…
8	0.0049	0.0006	0.0007	…
>8	0.0113	0.0007	0.0007	…

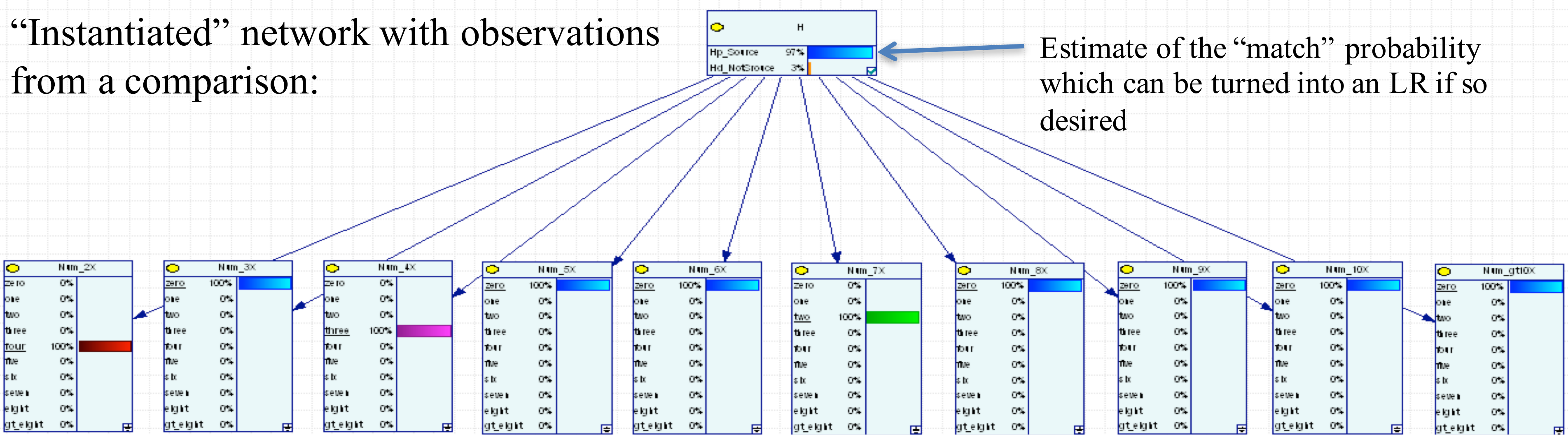
- So what can we use these for??
  - Lot’s of stuff, but we put them into a **Bayesian network**:
    - BN model for *Match/Non-match probabilities* given observed numbers of CMS runs

Bayesian Networks

"Prior" network based on Neel and Wells observed counts and Multinomial-Dirichlet model:



"Instantiated" network with observations from a comparison:



Estimate of the "match" probability which can be turned into an LR if so desired

Future Directions

- **Clean up:** cptID, feature2, fdrID
- **GUI modules** for common toolmark comparison tasks/calculations using 3D microscope data
- **2D features** for toolmark impressions
- **Parallel/GPU/FPGA** implementation of computationally intensive routines e.g. ALMA Correlator for astronomy data
  - Especially for retrieving “relevant pop/best match” reference sets
- **Uncertainty for Bayesian Networks**
  - Models, parameters…

References

Zheng: http://www.nist.gov/forensics/ballisticsdb/

Lillien:

- http://www.openfmc.org/
- http://www.cadreforensics.com/

Brubaker: https://github.com/OpenFMC/x3p

Petraco:

- https://github.com/npetraco/x3pr
- https://github.com/npetraco/feature2
- https://github.com/npetraco/cptID
- https://github.com/npetraco/fdrID

Edd:

- https://cran.r-project.org/web/packages/Rcpp/index.html
- http://dirk.eddelbuettel.com/code/rcpp.html

OpenCV: http://opencv.org/

Whitcher: https://cran.r-project.org/web/packages/waveslim/index.html

R-Core: https://www.r-project.org/

Vovk: http://www.alrw.net/

Efron:

- Efron, B. “Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction”, Cambridge, 2013.

References

locfdr: https://cran.r-project.org/web/packages/locfdr/index.html

JAGS:

- http://mcmc-jags.sourceforge.net/
- https://cran.r-project.org/web/packages/rjags/index.html

Stan:

- http://mc-stan.org/
- https://cran.r-project.org/web/packages/rstan/index.html

Neel:

- Neel, M and Wells M. "A Comprehensive Analysis of Striated Toolmark Examinations. Part 1: Comparing Known Matches to Known Non-Matches", AFTE J 39(3):176-198 2007.

Wevers:

- Wevers, G, Michael Neel, M and Buckleton, J. "A Comprehensive Statistical Analysis of Striated Tool Mark Examinations Part 2: Comparing Known Matches and Known Non-Matches using Likelihood Ratios", AFTE J 43(2):1-9 2011.

Buckleton:

- Buckleton J, Nichols R, Triggs C and Wevers G. "An Exploratory Bayesian Model for Firearm and Tool Mark Interpretation", AFTE J 37(4):352-359 2005.

BayesFusion: http://www.bayesfusion.com/

Acknowledgements

- Professor Chris Saunders (SDSU)
- Professor Christophe Champod (Lausanne)
- Alan Zheng (NIST)
- Ryan Lillien (Cadre)
- Scott Chumbley (Iowa State)
- Robert Thompson (NIST)
- Research Team:
  - Mr. Daniel Azevedo
  - Ms. Tatiana Batson
  - Dr. Martin Baiker
  - Ms. Julie Cohen
  - Dr. Peter Diaczuk
  - Mr. Antonio Del Valle
  - Ms. Carol Gambino
  - Dr. James Hamby
  - Mr. Nick Natalie
  - Mr. Mike Neel
  - Ms. Alison Hartwell, Esq.
  - Mr. Robert McLean
  - Dr. Brooke Kammrath
  - Mr. Chris Lucky
  - Off. Patrick McLaughlin
  - Dr. Mecki Prinz
  - Dr. Linton Mohammed
  - Ms. Diana Paredes
  - Mr. Nicholas Petraco
  - Ms. Stephanie Pollut
  - Dr. Jacqueline Speir
  - Dr. Peter Shenkin
  - Mr. Peter Tytell
  - Dr. Peter Zoon

Collaborations, Reprints/Preprints:

npetraco@gmail.com

http://jjcweb.jjay.cuny.edu/npetraco/