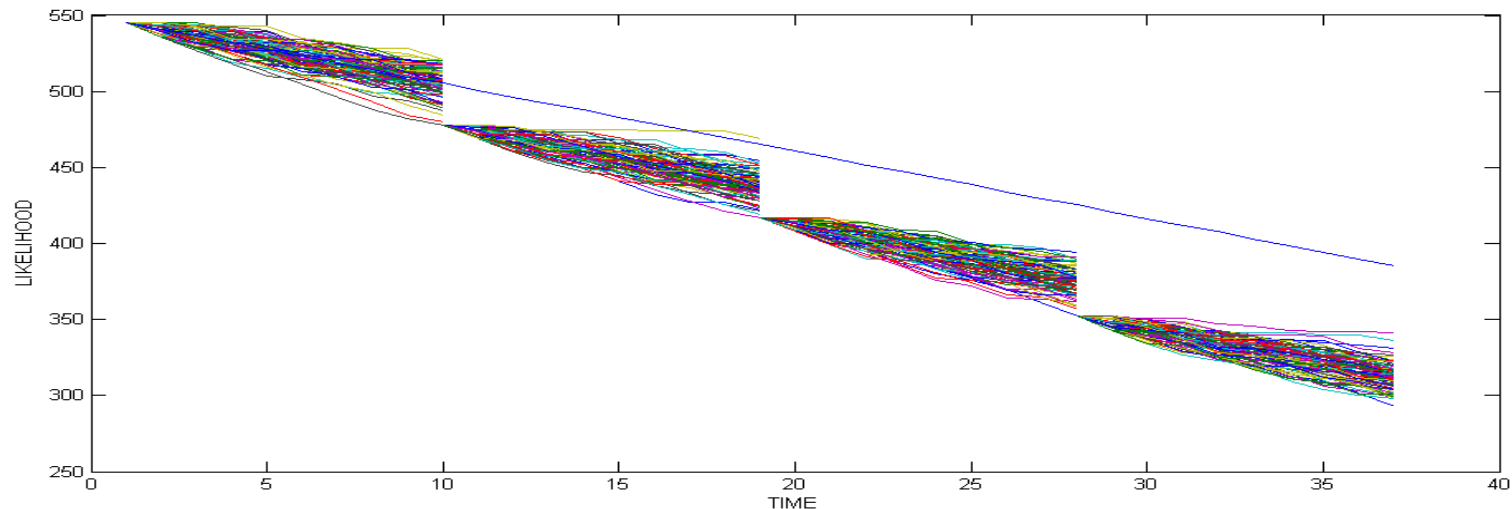


Research presentation by PhD Harri Pölönen

Experiences using Techila Grid in Quantification of Biomedical Data



M²oBSI

Methods and Models for Biological Signals and Images



- The 14 people M2oBSI research group is lead by professor **Ulla Ruotsalainen**
- The research area of the Methods and Models for Biological Signals and Images (M²oBSI) group is **biological and medical image reconstruction, processing and analysis**. The aim is to develop signal and image processing methods for automatic analysis of 3D functional images.



The Main Problem in the Research

- The whole research the M2oBSI group does relies heavily on computational science.
- Main problem is the availability of usable computing resources
- We prefer high abstraction language, like MATLAB, since we do constant iterative code development from scratch all the time instead of utilizing production ready code → plan, implement, execute, analyze, plan, implement, execute, analyze...
- Lower level languages like C/C++ would offer more computing speed but would also at the same time slow down code development speed → MATLAB is very suitable for efficient development purpose.



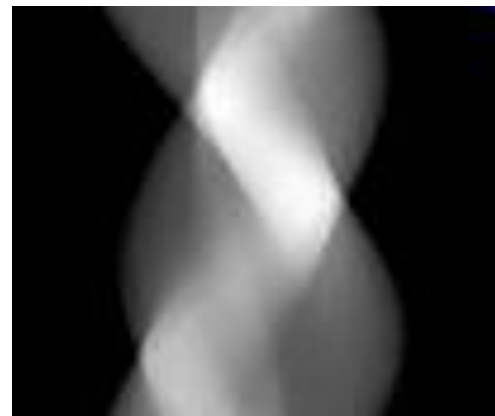
Positron Emission Tomography

- Maximum likelihood estimation of kinetic parameters from a time series of PET projections
- Huge number of parameters (up to 100 000) to be estimated simultaneously
- Large amount of data (up to tens of gigabytes)

PHANTOM



PET DATA



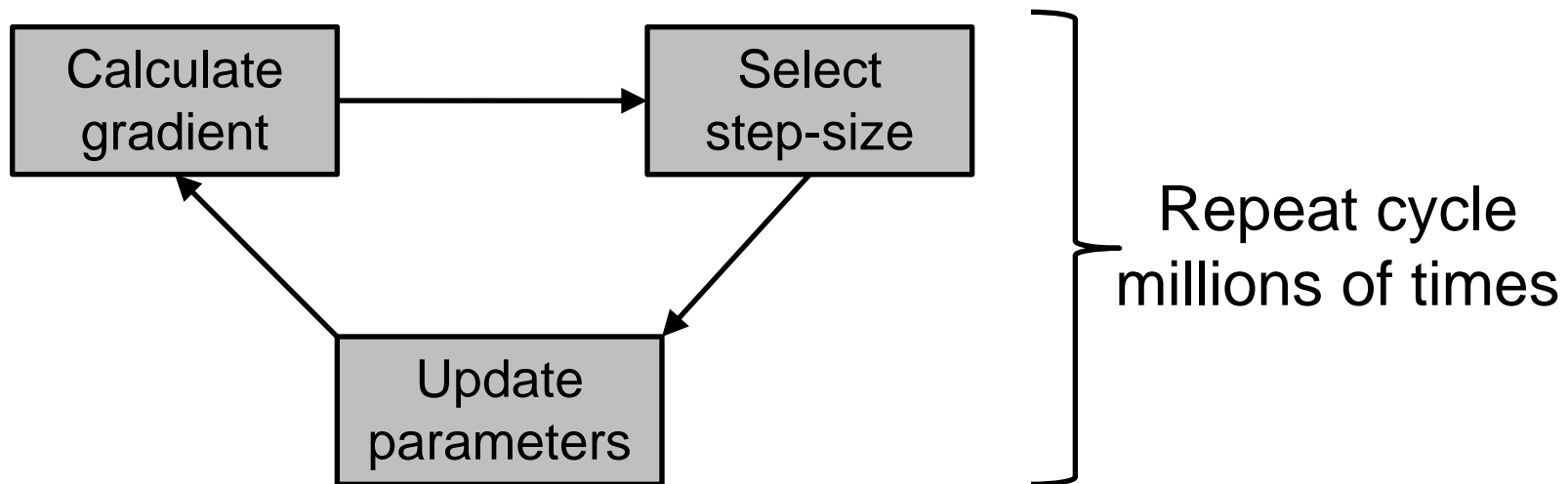
Harri Pölonen, Jari Niemi and Ulla Ruotsalainen:

Error-Corrected Estimation of Regional Kinetic Parameter Histograms Directly from PET Projections. Submitted to IOP Physics in Medicine and Biology.



Positron Emission Tomography

- The parameters can be estimated with a gradient method...
... but the optimal step-size towards gradient is unknown
- Barzilai-Borwein¹ has proved to be the best automatic step-size selection method
...but the estimation can still take several weeks on a single computer



¹J. Barzilai and J. M. Borwein, "Two-Point Step Size Gradient Methods," IMA J Numer Anal, vol. 8, no. 1, pp. 141–148, 1988.

Positron Emission Tomography

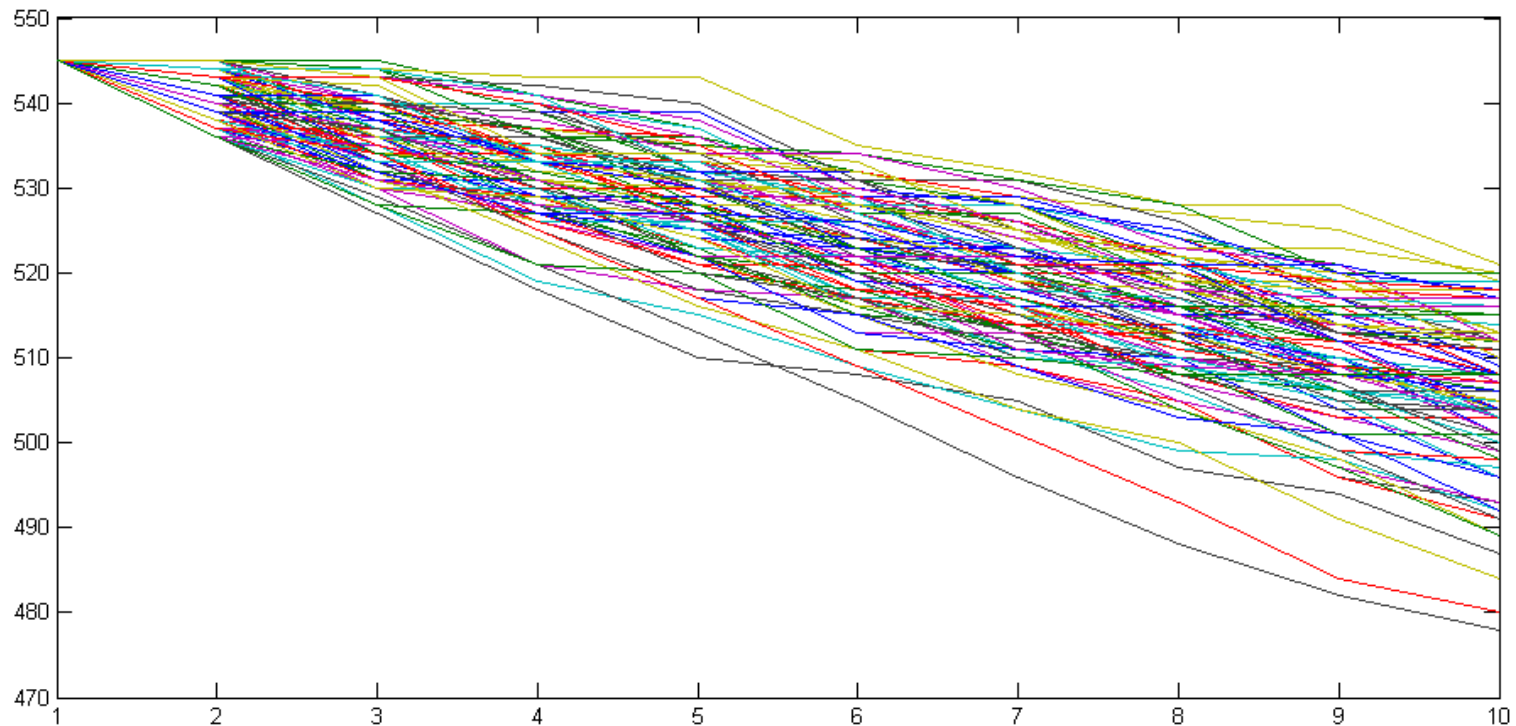
- The estimation problem can not be calculated in a grid directly, because all the parameters must be estimated simultaneously and are inter-dependent.
- But, let's take a non-conventional **guessing approach** in Techila grid...
 - A single guess for a gradient step-size may be good or bad
 - Within 1000 grid nodes there are lots of bad guesses...

... but also some **very good guesses!**



Positron Emission Tomography

Likelihood development of 50 nodes in 10 minutes (illustrative) with unique random step-sizes

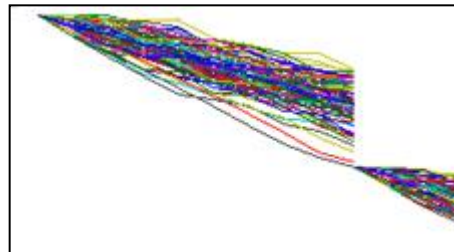


Note that due to good luck, some nodes show very good development!



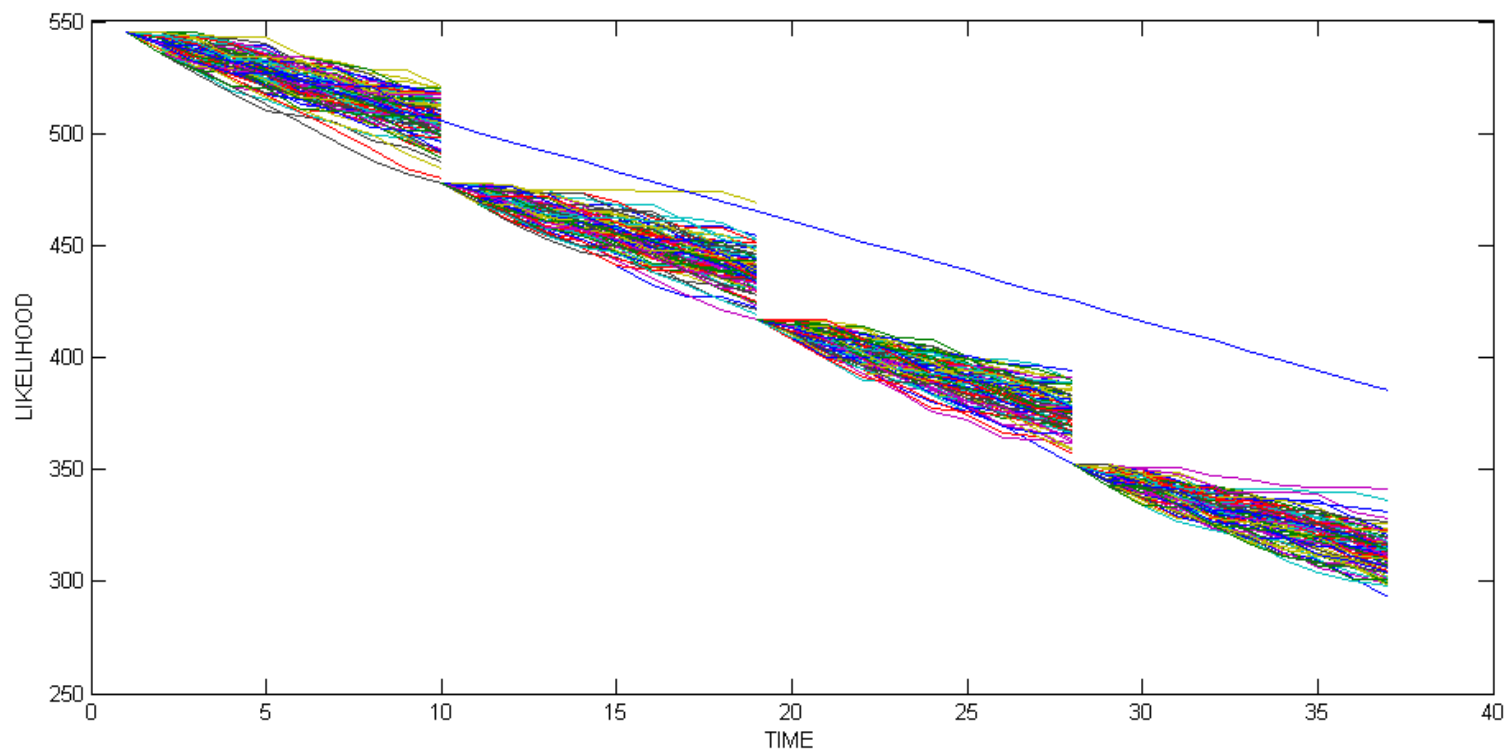
Positron Emission Tomography

- Over time luck settles and all nodes perform equally well
- With frequent intervals, let's initialize the nodes according to the best performing node
 - All the nodes continue from the current best guess solution



Positron Emission Tomography

Likelihood development with node re-initialization every 10 minutes
Blue line shows the standard Barzilai-Borwein method



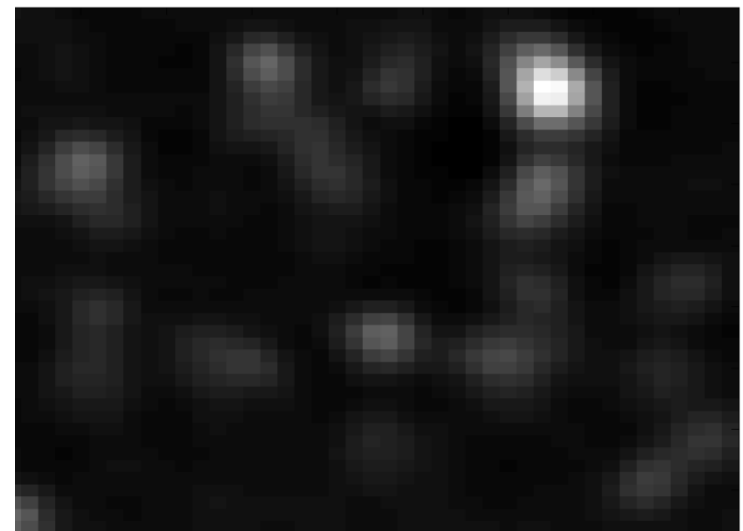
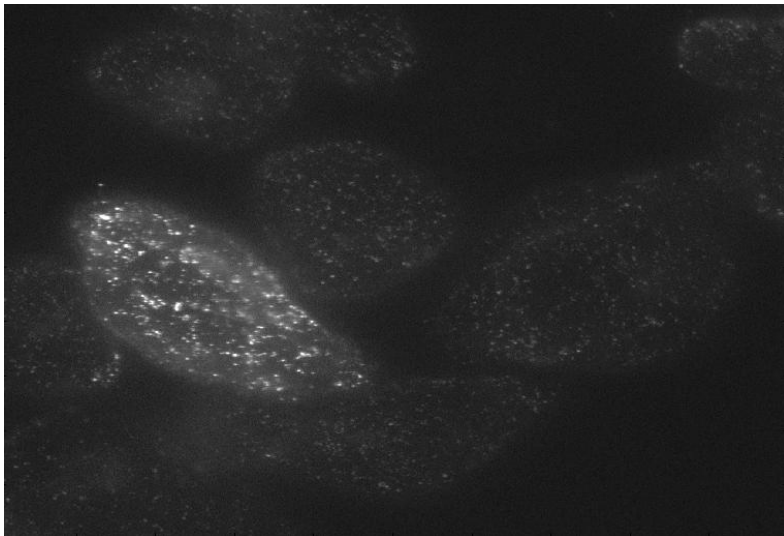
Positron Emission Tomography

- The more nodes we use -> better guesses are found -> faster optimization
 - Thus, we have improved a deterministic non-distributable problem through stochastic approach and clever use of large pool of underutilized IT-infra
 - **Estimates can now be found significantly faster by using the Techila Grid (in ~two days instead of ~two months)**
 - The approach relies on the large number of grid nodes rather than on the actual raw computational power
- This method benefits a lot from many nice features Techila Grid offers like time limits (forced to stop the computation after selected time limits), optimization (selects fastest free resources automatically), error tolerance, easy-to-use



Fluorescence microscopy data quantification

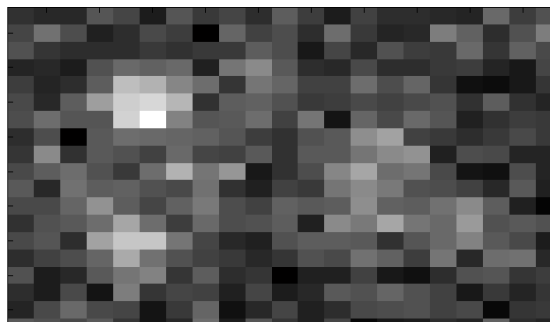
- The intensities, locations and shapes of spots in a microscope image are to be determined
- Each group of overlapping or closely located spots is quantified separately with an evolutionary algorithm



Fluorescence microscopy data quantification

- Each cell contained about 300 spots
- It took up to 15 minutes to process one cluster of spots on a single computer
- In total we had 60 cells in three treatment groups

Estimated time to complete on a single computer: **several months!**



Fluorescence microscopy data quantification

- Fortunately, the problems is directly separable to several sub-processes
→ embarrassingly parallel problem
- Each cluster of spots is quantified separately in Techila grid

Computation time drops from "several months" to ~10 minutes!

Maurice Jansen, Vilja M. Pietiäinen, Harri Pölonen, Laura Rasilainen, Mirkka Koivusalo, Ulla Ruotsalainen, Eija Jokitalo, and Elina Ikonen:
Cholesterol substitution increases the structural heterogeneity of caveolae. *J. Biol. Chem.*, Vol. 283, Issue 21, 14610-14618, May 23, 2008.

Harri Pölonen, Jussi Tohka and Ulla Ruotsalainen:

Automatic intensity quantification of fluorescence targets from microscope images with maximum likelihood estimation. *Proceeding of European Signal Processing Conference (EUSIPCO 2009), Glasgow, UK 2009.* (The article was chosen in the TOP 25 amongst 550 conference articles)

Harri Pölonen, Jussi Tohka and Ulla Ruotsalainen:

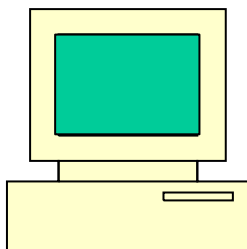
Automatic Quantification of Fluorescence from Clustered Targets in Microscope Images. *Lecture Notes in Computer Science*, Vol. 5575, p. 667-675, 2009.



Conclusion

- The huge number of nodes in the grid offers huge improvements to the computations time costs
- Some clever modifications may be needed to the methods, though

- Techila GRID proved to be an invaluable tool in Harri Pölönen's PhD thesis *Quantification of Biomedical Data with Stochastic Parametric Models and Numerical Optimization.*



1 vs. n

