

# Lorenz Wernisch

MRC Programme leader

## Contact

12 Chalmers Road  
Cambridge CB1 3SX  
+44 (0)1223 246690  
lwernisch@gmail.com

## Links

[www.linkedin.com/in/lwernisch](http://www.linkedin.com/in/lwernisch)  
[lwernisch.github.io](http://lwernisch.github.io)  
Google Scholar

## Research

### Statistics/ML

Bayesian statistics • Evidence synthesis • Gaussian processes • Bayesian nonparameters • MCMC • Variational inference • Time series • Active learning • Deep learning • Network inference • Graphical models • Dynamical Systems

### Biology/Medicine

Gene regulation • Protein structure • Stem cells • Single cells • Biomarkers • Intensive care data

### Computer science

Combinatorial optimisation • Algorithmic geometry • Graph theory

## Education

### Professional Accreditation of Teaching in Higher Education

2002 | University of London

### PhD in Mathematics

1990–1994 | Freie Universität Berlin  
Dominance order on point sets and aligned rectangles

### BSc/MSc in Mathematics

1982–1989 | Universität Wien  
Dimensions of posets and comparability graphs, with *Distinction*

### Study of Philosophy

1979–1982 | Universität Wien

## Skills

### Leadership

- 18 years experience in leading a research group (up to 6 members)
- 7 years experience as a MSc course (Bioinformatics) director
- 11 years leading a Medical Research Council programme of research

### Research

- Broad range of expertise in mathematical, computational, and statistical modelling, theoretical as well as applied
- Contributions to conceptual advances in Bayesian statistics
- 30+ years experience in algorithm and software development and implementation in a variety of frameworks (Bugs, Stan, Cuda, Tensorflow) and languages (Assembler, Lisp, C++, Java, R, Python) on various computational platforms (from early transputer to GPU systems)

### Training

- 25+ years experience in teaching mathematics, statistics, bioinformatics, computing
- For 7 years principal lecturer on an MSc course in Bioinformatics, Birkbeck College, London, now contributing to the MPhil Computational Biology course at Cambridge University
- Supervision of 12 successful PhD students and countless MSc students

### Management

- Grant management as principal investigator
- Collaborative research projects with biologists, medical researchers and clinicians

## Experience

### Programme leader, statistical genomics

since 4/2007 | Medical Research Council, Biostatistics Unit | Cambridge  
Leading a programme on computational biology and statistical genomics.

### Lecturer/Reader/Professor in Computational Biology

10/1999–3/2007 | Birkbeck College | University of London

Developing research programmes on protein structure analysis, microarray analysis, gene regulation and network inference.

### Research scientist

10/1996–9/1999 | EMBL European Bioinformatics Institute | Cambridge

Developing algorithms and software for the analysis of protein structure interfaces and optimisation of peptide binding.

### Research assistant and postdoctoral fellow

5/1990–9/1996 | Department of mathematics, Freie Universität Berlin

Working on algorithmic geometry and discrete optimisation problems. Teaching on undergraduate computer science courses.

### Software developer

4/1987–4/1990 | Elin ALCATEL-Austria Research Center Wien

Lead developer of a system acquiring test data from electric motors, later contributing to the development of knowledge based inference engine on parallel transputer systems.

## Current research themes

---

### Evidence synthesis

There is growing interest in analysing large heterogeneous data sets jointly, that provide different views on common underlying features. We have recently proposed a principled framework to merge separate Bayesian models, each tailored to a specific data set, into one model through common parameters (Goudie et al., 2018). A key requirement is the ability to estimate densities. We are exploring the potential of deep learning approaches, such as normalising flows, to provide detailed density estimates even for sparse data in high dimensions and the integration of such estimations into Monte Carlo sampling from or variational optimisation of Bayesian models. Another aspect of integrating heterogeneous data is the restriction of the flow of information through a so called 'cut' in a Bayesian model. There has been much debate about its meaning and we are developing a principled framework for such operator based on variational inference. The ability to estimate densities reliably plays a key role in this approach.

### Design of experiments using active learning

Controlled differentiation of stem cells is driven by careful administration of stimulating factors to cell cultures. The exact amount and timing of such stimulation is largely a matter of trial and error. We propose an active learning framework implemented in Tensorflow to optimise such dynamic experiments. Since, unlike the typical situation for many tasks in reinforcement learning, the number of possible experiments is extremely limited, modelling of the dynamics through Gaussian processes state space models (instead of a vanilla neural networks) is crucial. In simulations of realistic systems we achieve remarkable optimisation results with only a limited number of experiments (see [this talk at Costnet18](#)). Some earlier work on active learning for gene networks is described in Pournara and Wernisch, 2004.

### Prediction in intensive care

With the increasing availability of health care data, including measurements of biomarkers, there is growing interest in the prediction and the understanding of adverse events such as a secondary infection of patients in intensive care. Modeling of trajectories as well as the dynamics through Gaussian processes is a promising route we apply to the prediction of secondary infections as well as the long-term effect of traumatic brain injury. A key step is the understanding of the causal relationship of components of the immune response. We explore how far concepts of NLTSA (non-linear time series analysis) such as convergent cross mapping (Sugihara et al. 201)

can be put on a firm probabilistic basis and used for this purpose. Details can be found in a tutorial on ML for ICU data.

### Nonparametric approaches to single cell data

Transcriptomics data from single cells pose a unique challenge in that, although measurements from thousands of samples are available knowledge of their temporal ordering is typically very limited. In a series of studies (Strauß et al., 2018, Gabasova et al., 2017, Reid and Wernisch, 2016, Penfold et al., 2018) we derive pseudo-time orderings and clusterings using Gaussian processes and Dirichlet process priors. We are now exploring how far such reconstructed gene expression trajectories can be exploited for network inference (some examples given in [a talk on single cell analysis](#)). We have also expanded the approach of Gabasova et al., 2017, to a generic framework for clustering using hierarchical Dirichlet distributions.

*For a full list of publications see [Google Scholar](#), for a list with links see [lwernisch.github.io/publications](#)*

## Recent invited talks

---

*Gene regulatory networks from single cell data, SMPGD2017, Statistical Methods for Postgenomic Data, Imperial College, 2017*

*Improving protocols for cell differentiation through reinforcement learning, Workshop Statistical Network Science, Brunel University, 2018*

*Improving protocols for cell differentiation through reinforcement learning, Costnet2018, Warsaw, 2018*

## Grants

---

- *Computational inference of optimal routes for microarray experiments*, 2003–2006, BBSRC, 160K GBP
- *Functional Genomics of Neuronal Cells*, 2002–2007, Wellcome Trust Functional Genomics Programme Grant with A Lumsden (King's College), G Koentges (UCL), T Jessell (Columbia, US), and C Dulac (Harvard, US), 715K GBP (of 5.4M GBP total)
- *Interdisciplinary Programme for Cellular Regulation*, EPSRC Multidisciplinary Critical Mass grant, 2004–2007, EPSRC collaborative grant with Prof D Rand (Univ. of Warwick) and others

## Other interests

---

Violin • String quartet • Photography • Video • EDM production • Flying (PPL, aerobatics) • Volleyball • Windsurfing • Skiing

## Selected publications

Robert JB Goudie, Anne M Presanis, David Lunn, Daniela De Angelis, and Lorenz Wernisch. Joining and splitting models with markov melding. *Bayesian Analysis*, 2018.

Christopher A Penfold, Anastasiya Sybirna, John E Reid, Yun Huang, Lorenz Wernisch, Zoubin Ghahramani, Murray Grant, and M Azim Surani. Branch-recombinant Gaussian processes for analysis of perturbations in biological time series. *Bioinformatics*, 34(17):i1005–i1013, September 2018.

Magdalena E Strauß, John E Reid, and Lorenz Wernisch. GPseudoRank: a permutation sampler for single cell orderings. *Bioinformatics*, July 2018.

Felicia S L Ng, David Ruau, Lorenz Wernisch, and Berthold Göttgens. A graphical model approach visualizes regulatory relationships between genome-wide transcription factor binding profiles. *Brief. Bioinformatics*, 19(1):162–173, 01 2018.

Evelina Gabasova, John Reid, and Lorenz Wernisch. Clusternomics: Integrative context-dependent clustering for heterogeneous datasets. *PLoS Comput. Biol.*, 13(10):e1005781, October 2017.

Richard Newton and Lorenz Wernisch. A comparison of machine learning and Bayesian modelling for molecular serotyping. *BMC Genomics*, 18(1):606, August 2017.

John E Reid and Lorenz Wernisch. Pseudotime estimation: deconfounding single cell time series. *Bioinformatics*, 32(19):2973–80, 10 2016.

Deepak Parashar, Jack Bowden, Colin Starr, Lorenz Wernisch, and Adrian Mander. An optimal stratified Simon two-stage design. *Pharm Stat*, 15(4):333–40, 07 2016.

Richard Newton and Lorenz Wernisch. Investigating inter-chromosomal regulatory relationships through a comprehensive meta-analysis of matched copy number and transcriptomics data sets. *BMC Genomics*, 16:967, November 2015.

Richard Newton and Lorenz Wernisch. A meta-analysis of multiple matched copy number and transcriptomics data sets for inferring gene regulatory relationships. *PLoS ONE*, 9(8):e105522, 2014.

Fernando J Calero-Nieto, Felicia S Ng, Nicola K Wilson, Rebecca Hannah, Victoria Moignard, Ana I Leal-Cervantes, Isabel Jimenez-Madrid, Evangelia Diamanti, Lorenz Wernisch, and Berthold Göttgens. Key regulators

control distinct transcriptional programmes in blood progenitor and mast cells. *EMBO J.*, 33(11):1212–26, June 2014.

John E Reid and Lorenz Wernisch. STEME: a robust, accurate motif finder for large data sets. *PLoS ONE*, 9(3):e90735, 2014.

Marina Evangelou, Frank Dudbridge, and Lorenz Wernisch. Two novel pathway analysis methods based on a hierarchical model. *Bioinformatics*, 30(5):690–7, March 2014.

Xin Yi Goh, Richard Newton, Lorenz Wernisch, and Rebecca Fitzgerald. Testing the utility of an integrated analysis of copy number and transcriptomics datasets for inferring gene regulatory relationships. *PLoS ONE*, 8(5):e63780, 2013.

Dennis Wang, Augusto Rendon, and Lorenz Wernisch. Transcription factor and chromatin features predict genes associated with eQTLs. *Nucleic Acids Res.*, 41(3):1450–63, February 2013.

Yang Luo, Chea Lu Lim, Jennifer Nichols, Alfonso Martinez-Arias, and Lorenz Wernisch. Cell signalling regulates dynamics of Nanog distribution in embryonic stem cell populations. *J R Soc Interface*, 10(78):20120525, January 2013.

Marina Evangelou, Augusto Rendon, Willem H Ouwehand, Lorenz Wernisch, and Frank Dudbridge. Comparison of methods for competitive tests of pathway analysis. *PLoS ONE*, 7(7):e41018, 2012.

Dennis Wang, Augusto Rendon, Willem Ouwehand, and Lorenz Wernisch. Transcription factor co-localization patterns affect human cell type-specific gene expression. *BMC Genomics*, 13:263, June 2012.

John E Reid and Lorenz Wernisch. STEME: efficient EM to find motifs in large data sets. *Nucleic Acids Res.*, 39(18):e126, October 2011.

Richard Newton, Jason Hinds, and Lorenz Wernisch. Empirical Bayesian models for analysing molecular serotyping microarrays. *BMC Bioinformatics*, 12:88, March 2011.

Yi Zhang, Kim A Hatch, Joanna Bacon, and Lorenz Wernisch. An integrated machine learning approach for predicting DosR-regulated genes in *Mycobacterium tuberculosis*. *BMC Syst Biol*, 4:37, March 2010.

Núria Domedel-Puig, Iosifina Pournara, and Lorenz Wernisch. Statistical model comparison applied to common network motifs. *BMC Syst Biol*, 4:18, March 2010.

- John E Reid, Kenneth J Evans, Nigel Dyer, Lorenz Wernisch, and Sascha Ott. Variable structure motifs for transcription factor binding sites. *BMC Genomics*, 11:30, January 2010.
- John E Reid, Sascha Ott, and Lorenz Wernisch. Transcriptional programs: modelling higher order structure in transcriptional control. *BMC Bioinformatics*, 10:218, July 2009.
- Mario dos Reis and Lorenz Wernisch. Estimating translational selection in eukaryotic genomes. *Mol. Biol. Evol.*, 26(2):451–61, February 2009.
- Iosifina Pournara and Lorenz Wernisch. Using temporal correlation in factor analysis for reconstructing transcription factor activities. *EURASIP J Bioinform Syst Biol*, page 172840, 2008.
- Yi Zhang, Kim A Hatch, Lorenz Wernisch, and Joanna Bacon. A Bayesian Change point model for differential gene expression patterns of the DosR regulon of *Mycobacterium tuberculosis*. *BMC Genomics*, 9:87, February 2008.
- Kenneth Evans, Sascha Ott, Annika Hansen, Georgy Koentges, and Lorenz Wernisch. A comparative study of S/MAR prediction tools. *BMC Bioinformatics*, 8:71, March 2007.
- Iosifina Pournara and Lorenz Wernisch. Factor analysis for gene regulatory networks and transcription factor activity profiles. *BMC Bioinformatics*, 8:61, February 2007.
- Richard Newton, Jason Hinds, and Lorenz Wernisch. A Hidden Markov model web application for analysing bacterial genotyping DNA microarray experiments. *Appl. Bioinformatics*, 5(4):211–8, 2006.
- Mike Withers, Lorenz Wernisch, and Mario dos Reis. Archaeology and evolution of transfer RNA genes in the *Escherichia coli* genome. *RNA*, 12(6):933–42, June 2006.
- Núria Domedel-Puig and Lorenz Wernisch. Applying GIFT, a Gene Interactions Finder in Text, to fly literature. *Bioinformatics*, 21(17):3582–3, September 2005.
- Martin T Edwards, Stuart C G Rison, Neil G Stoker, and Lorenz Wernisch. A universally applicable method of operon map prediction on minimally annotated genomes using conserved genomic context. *Nucleic Acids Res.*, 33(10):3253–62, 2005.
- Mario dos Reis, Renos Savva, and Lorenz Wernisch. Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Res.*, 32(17):5036–44, 2004.
- Iosifina Pournara and Lorenz Wernisch. Reconstruction of gene networks using Bayesian learning and manipulation experiments. *Bioinformatics*, 20(17):2934–42, November 2004.
- Mario dos Reis, Lorenz Wernisch, and Renos Savva. Unexpected correlations between gene expression and codon usage bias from microarray data for the whole *Escherichia coli* K-12 genome. *Nucleic Acids Res.*, 31(23):6976–85, December 2003.
- Lorenz Wernisch and Shoshana J Wodak. Identifying structural domains in proteins. *Methods Biochem Anal*, 44:365–85, 2003.
- Lorenz Wernisch, Sharon L Kendall, Shamit Soneji, Andreas Wietzorrek, Tanya Parish, Jason Hinds, Philip D Butcher, and Neil G Stoker. Analysis of whole-genome microarray replicates using mixed models. *Bioinformatics*, 19(1):53–61, January 2003.
- Alfonso Jaramillo, Lorenz Wernisch, Stéphanie Héry, and Shoshana J Wodak. Folding free energy function selects native-like protein sequences in the core but not on the surface. *Proc. Natl. Acad. Sci. U.S.A.*, 99(21):13554–9, October 2002.
- Lorenz Wernisch, Stéphanie Hery, and Shoshana J Wodak. Automatic protein design with all atom force-fields by exact and heuristic optimization. *Journal of molecular biology*, 301(3):713–736, 2000.
- Lorenz Wernisch, Marcel Hunting, and Shoshana J Wodak. Identification of structural domains in proteins by a graph heuristic. *Proteins: Structure, Function, and Bioinformatics*, 35(3):338–352, 1999.
- Stefan Felsner, Rudolf Müller, and Lorenz Wernisch. Trapezoid graphs and generalizations, geometry and algorithms. *Discrete Applied Mathematics*, 74(1):13–32, 1997.
- Jiří Matoušek, Emo Welzl, and Lorenz Wernisch. Discrepancy and approximations for bounded vc-dimension. *Combinatorica*, 13(4):455–466, 1993.
- Stefan Felsner and Lorenz Wernisch. Maximum k-chains in planar point sets: Combinatorial structure and algorithms. *SIAM Journal on Computing*, 28(1):192–209, 1998.
- Stefan Felsner, Rudolf Müller, and Lorenz Wernisch. Trapezoid graphs and generalizations, geometry and algorithms. In *Scandinavian Workshop on Algorithm Theory*, pages 143–154. Springer, Berlin, Heidelberg, 1994.
- Stefan Felsner and Lorenz Wernisch. Markov chains for linear extensions, the two-dimensional case. In *SODA*, pages 239–247, 1997.