34th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering

21-26 September 2014

Château Clos Lucé, Amboise - France

«Mathematics embrace all the things of the universe»

Leonardo **da Vinci**

MaxEnt 2014 SPONSORS

Scientific sponsor

Société Mathématique de France

SMF

SMAI

Funding sponsors

BOISE STATE UNIVERSITY

entropy

Logistic sponsors

Supélec

UNIVERSITÉ PARIS SUD

Comprendre le monde, construire l'avenir®

cnrs

see

Welcome to MaxEnt 2014 Conference

On behalf of both organizing and scientific committees, it is a great pleasure to welcome all delegates, representatives and participants from around the world to MAXENT 2014 conference. The thirty-Fourth International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering will be held in France organized by SEE (https://www.see.asso.fr/en) under the auspices of « Centre national de la recherche scientifique (CNRS) », « École supérieure d'électricité (SUPÉLEC) » and « Université Paris Sud (UPS), Orsay». MAXENT 2014 also benefits from scientific support of SMF (Société Mathématique de France: http://smf.emath.fr/), and financial sponsoring from Jaynes foundation.

Proceedings of MaxEnt 2014 will be published by AIP (American Institute of Physics).

MDPI Entropy publication is an editorial partner to publish long papers issue from MaXent 2014 Tutorial, Invited and selected authors in special issue on "Information, Entropy and their Geometric Structures". All other interested authors are invited to submit their long papers in this Special Issue:

http://www.mdpi.com/journal/entropy/special_issues/entropy-Geome.

MaxEnt 2014 will take place in Château Le Clos Lucé (http://www.vinci-closluce.com/en/) the last residence of Leonardo Da Vinci, in Amboise at the heart of Loire Valley Castles, labeled by UNESCO as World heritage. Le Clos Lucé is at proximity of King Francis the First Amboise Castle, apex of the French Renaissance and of Vouvray wines vineyards and cellars. Let us express all our thanks to the Clos Lucé Castle and Saint Bris Family for hosting this event in this historical place.

MaxEnt 2014 strives to present Bayesian inference and Maximum Entropy methods in data analysis, information processing and inverse problems from a broad range of diverse disciplines: Astronomy and Astrophysics, Geophysics, Medical Imaging, Molecular Imaging and genomics, Non Destructive Evaluation, Particle and Quantum Physics, Physical and Chemical Measurement Techniques, Economics and Econometrics.

This year special interest will be **Geometrical Sciences of Information / Information Geometry and their link with classical subjects of MaxEnt workshops which are Entropy, Maximum Entropy and Bayesian inference in sciences and engineering**. The focus will be more on using these concepts in generic Inverse problems, multidimensional and multi components Time Series Analysis and Spectral Estimation, Deconvolution and Source Separation, Segmentation, Classification and Pattern Recognition, X-ray, Diffractive, Diffusive and Quantum Tomographic Imaging.

Specifically this year, MaxEnt 2014 technical program is covering topics and highlights in the domain of “Geometric Science of Information” including Information Geometry Manifolds of structured data/information and their advanced applications. These topics addresses inter-relations between different mathematical domains like shape spaces (geometric statistics on manifolds and Lie groups, deformations in shape space,…), probability/optimization & algorithms on manifolds (structured matrix manifold, structured data/Information, …), relational and discrete metric spaces (graph metrics, distance geometry, relational analysis,…), computational and hessian information geometry, algebraic/infinite dimensionnal/Banach information manifolds, divergence geometry, tensor-valued morphology, optimal transport theory, manifold & topology learning, ….

Papers are presented in Tutorials, Keynotes, Oral and Poster sessions. Publications will be presented and testify the world wide interest for topics covered by MaxEnt. After **Sunday Tutorials:**

Modern Probability Theory by Kevin H. Knuth

The Basics of Information Geometry by Ariel Caticha

Voronoi diagrams in information geometry by Franck Nielsen

Foundations and Geometry by John Skilling

Uncertainty quantification for computer model by Udo V. Toussaint

Geometric Structures Induced From Divergence Functions by Jun Zhang

Koszul Information Geometry and Souriau Lie Group Thermodynamics by Frédéric Barbaresco

Bayesian and Information Geometry in signal processing by Ali Mohammad-Djafari,

Each day, we will have keynote presentations by international experts and in particular this year we have the following invited:

- **“On the Structure of Entropy”** by Prof. **Mikhail GROMOV** (Abel Prize, IHES, Bures-sur-Yvette, France)

- **“Topological forms of information”** by Prof. **Daniel BENNEQUIN** (member of l'Institut Mathématique de Jussieu, Denis Diderot University, Paris, France)

- “***The entropy-based quantum metric***” by Prof. **Roger BALIAN** (Member of French Academy of Sciences, Scientific Consultant of CEA, France)

- “***Duhem’s abstract thermodynamics***” by Prof. **Stefano BORDONI** (associate professor of logic, philosophy and history of science, Bologna University, Italy)

We would like to acknowledge all the Organizing and Scientific Committee members for their hard work, in evaluating submissions. We also give our thanks to authors and co-authors, for their tremendous effort and scientific contribution.

General chairs of MaxEnt 2014

Ali Mohammad-Djafari

Research Director at CNRS

Laboratoire des signaux et systèmes (L2S)

CNRS-SUPELEC-UNIV PARIS SUD

3 rue Joliot-Curie, 91192 Gif-sur-Yvette

Frédéric BARBARESCO

SEE/SMF MaxEnt General Chair

President of SEE Club ISIC

« Ingéniérie des Systèmes d'Information

et de Communications »

(https://www.see.asso.fr/ct-isic)

SEE Emeritus Member, Ampere Medal 2007

Aymé Poirson Prize 2014 from French Academy of Sciences

Senior Scientist & Advanced Studies Manager, Thales Air Systems

ABOUT SEE

http://www.see.asso.fr

SEE (Société de l'Electricité, de l'Electronique et des TIC – Society for Electricity, Electronics and ICT) is a nonprofit scientific and technical organization mainly active in France and by extension in French-speaking countries. It aims at gathering and animating a community of persons and organizations concerned by Science and Technologies in the fields of Energy, Electronics and Communications to foster the progress of both theoretical approaches in these fields and new applications in the main sectors of Economy (Energetic transition, new Transportation challenges, Health and silver Economy, Digital life, etc.).

The two main vectors of action of SEE are the organization of S&T events (like MAXENT) and the publication of 2 periodic reviews (REE and 3EI). SEE has strong actions shared with “sister” Societies like IEEE.

Alain BRENAC

SEE General Secretary

SEE Emeritus Member

MaxEnt 2014 Conference

For more than 30 years the MaxEnt workshops have explored the use of Bayesian and Maximum Entropy methods in scientific and engineering applications. All aspects of probabilistic inference, such as foundations, techniques, links with physics and applications in sciences and engineering as well as in social and life science, are of interest.

34th International Workshop on Bayesian Inference and Maximum Entropy

Methods in Science and Engineering

21-26 September 2014

Château Clos Lucé, Parc Leonardo Da Vinci, Amboise, France

Invited talks are scheduled with 4 invited and tutorial speakers:

Prof. Mikhail Gromov (Abel Prize, IHES, Bures-sur-Yvette, France)
On the Structure of Entropy

Prof. Daniel Bennequin (Denis Diderot University, Paris, France)
Topological forms of information

Prof. Roger Balian (French Academy of Sciences Member, Scientific Consultant of CEA)
The entropy-based quantum metric

Dr. Stefano Bordoni (Bologna University, Italy)
Duhem’s abstract thermodynamics.

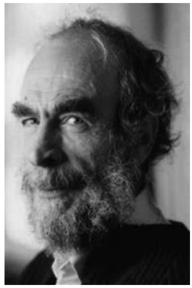
MaxEnt 2014 Organization

Program Committee :

Jean-François Bercher	ESIEE, France
Pierre Bessière	CNRS-Collège de France
Eric Barat	CEA, Saclay, France
Arshia Cont	IRCAM, France
Christophe Vignat	Université Paris Sud, France
Robert Niven	University of New South Wales, Australia
Adom Giffin	Clarkson University, NY, USA
Philip Goyal	pgoyal@albany.edu
Simone Fiori	Università Politecnica delle Marche, Italy
Vincent Vigneron	Universite d'Evry, France
John Skilling	Maximum Entropy Data Consultants
Richard Nock	Université des Antilles et de la Guyane,
Larrry Bretthorst	Washington Univ., USA
Ariel Caticha	University at Albany (SUNY), USA
Frank Nielsen	Sony Computer Science Lab., Inc., Tokyo, JAPAN
Barry Stocke	NewCastel, Australia
Rainer Fischer	Max-Planck-Institut, Germany
Paul Goggans	University of Mississippi, USA
Kevin Knuth	University at Albany (SUNY), USA
Frederic Barbaresco	SEE and Thales Air Systems, France
Ali Mohammad-Djafari	LSS-CNRS, France
Carlos Rodriguez	University at Albany (SUNY), USA
John Skilling	Cambridge, UK
Hichem Snoussi	Université de Technologie de Troyes
Jun Zhang	Univ. of Michigan, USA
Udo von Toussaint	Institut fuer Plasmaphysik, Germany

MaxEnt 2014 invited and tutorial speakers

Invited talks are scheduled with 4 invited and tutorial speakers:



Prof. Mikhail Gromov (Abel Prize, IHES,

Bures-sur-Yvette, France)

<u>On the Structure of Entropy</u>

Abstact: Mathematics is about "interesting structures". What make a structure interesting is an abundance of interesting problems; we study a structure by solving these problems. The worlds of science, as well as of mathematics itself, is abundant with gems (germs?) of simple beautiful ideas. When and how may such an idea direct you toward beautiful mathematics? I present in this talk a 20th century mathematician's perspective on Boltzmann's idea of entropy.

Biography: Mikhail Leonidovich Gromov. a French–Russian mathematician. He studied in the Leningrad University where he was a student of Vladimir Rokhlin. He has been working on the Riemannian Geometry, Geometric PDE, Symplectic Geometry, Geometric Theory of Groups and also on mathematical formalizations of ideas coming from biology, psychology and linguistic.

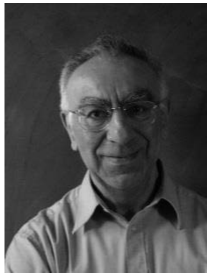
Prof. Daniel Bennequin

(Denis Diderot University, Paris, France)

<u>Topological forms of information</u>

Abstract: This talk will present recent joint works with Pierre Baudot, where we propose a general definition of categories of informations, and study a natural cohomology for associated information quantities, and homotopical derived notions. Entropies of Shannon, Kullback and Von Neumann, appear as first fundamental classes for classical and quantum setting respectively. The decomposition of entropy in higher mutual information functions appears as an homotopical structure, and generates a new kind of topology. Possible applications to the study of large statistical data and dynamics of neuronal systems will be mentioned

Biography: Born the 3 January 1952. Graduated from Ecole Normale Supérieure, he has defended his PhD in 1982 with Alain Chenciner at Paris VII University (Doctorat d'Etat, Entrelacements et équations de Pfaff). He was Professor at Strasbourg University and was member of Bourbaki. He is currently Professor at Denis Diderot University and member of l'Institut Mathématique de Jussieu. He has made many major contributions to contact geometry during 80's and was initiator of Contact Topology with Yakov Eliashberg. During the years 1990, he worked, with his students and colleagues in Strasbourg, on integrable systems and geometrical structures of Mathematical Physics. Since 2000 he is working in Neuroscience (mainly in LPPA directed by A.Berthoz, C-d-F, Paris); he made contributions to the study of geometrical invariance in human movements duration, dynamical structure of vestibular end sensors, organization of vestibular information flow, eye movements preparation, and gaze functions during locomotion.



Prof. Roger Balian

(French Academy of Sciences Member, Scientific Consultant of CEA)

The entropy-based quantum metric

Abstract: The von Neumann entropy S(^D) generates in the space of quantum density matrices ^D the Riemannian metric ds² = - d²S(^D), which is physically founded and which characterizes the amount of quantum information lost by mixing ^D and ^D + d^D. A rich geometric structure is thereby implemented in quantum mechanics. It includes a canonical mapping between the spaces of states and of observables, which involves the Legendre transform of S(^D). The Kubo scalar product is recovered within the space of observables. Applications are given to equilibrium and non-equilibrium quantum statistical mechanics. There the formalism is specialized to the relevant space of observables and to the associated reduced states issued from the maximum entropy criterion, which result from the exact states through an orthogonal projection. Von Neumann's entropy specializes into a relevant entropy. Comparison is made with other metrics. The Riemannian properties of the metric ds² = - d²S(^D) are derived. The curvature arises from the non-Abelian nature of quantum mechanics; its general expression and its explicit form for q-bits are given.

Biography: Roger Balian has been working at the "Institut de Physique Théorique" of Saclay (CEA), which he has directed (1979-1987). He was also Professor of Statistical Physics at Ecole Polytechnique (1972-1998), and Director of the "Ecole d'Eté de Physique Théorique des Houches" (1972-1980). His research works have addressed various topics, often related to statistical physics: superfluid Helium 3, signal theory, information/entropy, waves and complex trajectories, foundations of quantum mechanics, Casimir effect, quantum liquids, nuclear structure, gauge theories, distribution of galaxies...

Dr. Stefano Bordoni (Bologna University, Italy)

Duhem's abstract thermodynamics.

Abstract: In the second half of the nineteenth century, two different traditions of research emerged from Clausius' thermodynamics. Maxwell and Boltzmann pursued the integration of thermodynamics with the kinetic theory of gases, whereas others relied on a macroscopic and more abstract approach, which set aside specific mechanical models. Massieu, Gibbs, and Helmholtz exploited the structural analogy between mechanics and thermodynamics, the young Planck and J.J. Thomson aimed at filling the gap between thermodynamics and the theory of elasticity, and Oettingen developed a dual mathematical structure for heat and work. Starting from 1891, Pierre Duhem put forward the most original and systematic reinterpretation of abstract thermodynamics, and at the same time the boldest upgrade of Analytical mechanics. He developed and transformed the second tradition: his design of a generalized mechanics based on thermodynamics led to an astonishing mathematical unification between physics and chemistry. Purely mechanical phenomena and chemical reactions represented the opposite poles in Duhem's *Energetics*.

Biography: He graduated in physics, and received a PhD in History of science and then a PhD in Philosophy. He has recently gained a qualification as associate professor of Logic, philosophy and history of science. He has published papers and books on the history of science, and in particular history of physics. He has given seminars and lectures in some Italian universities and at the Max-Planck-Institut für Wissenschaftsgeschichte in Berlin. He has lectured in History of physics, History of science, and Mathematics.

MaxEnt 2014 SPECIAL SESSIONS & POSTERS

Sunday, September 21st

09:00 - 10:00 Registration

10:00 – 13:00 Tutorial session 1

Chaired by Ali Mohammad-Djafari / Frédéric Barbaresco / John Skilling

1. Modern Probability Theory (71) - Kevin H. Knuth
2. The Basics of Information Geometry (85) - Ariel Caticha
3. Voronoi diagrams in information geometry (94) - Franck Nielsen

13:00 - 14:30 Lunch Breack

14:30 – 19:30 Tutorial session 2

Chaired by Ariel Caticha / Ken H. Knuth / Frank Nielsen

4. Foundations and Geometry (104) - John Skilling
5. Uncertainty quantification for computer model (101) - Udo V. Toussaint
6. Geometric Structures Induced From Divergence Functions (95) - Jun Zhang
7. Koszul Information Geometry and Souriau Lie Group Thermodynamics (9)
   Frédéric Barbaresco
8. Bayesian and Information Geometry in signal processing (1)
   Ali Mohammad-Djafari

19:30 – 21:30 Welcome Cocktail

Monday, September 22nd

08:00 - 09:00 Registration

09:00 - 09:15 Opening Session

A. Mohammad-Djafari & F. Barbaresco

09:15 - 11:15 Oral session 1: Statistical Manifolds

Chaired by: Steeve Zozor / J.F. Bercher / Ariel Caticha /

67 Jean-Daniel Boissonnat, David Cohen-Steiner, Pooran Memari and Arnaud Poinas, ***Triangulating Statistical Manifolds***

69 Nina Miolane and Xavier Pennec, ***Statistics on Lie groups : a need to go beyond the pseudo-Riemannian framework***

76 Max H. M. Costa, Chandra Nair and Olivier Rioul, ***From Almost Gaussian to Gaussian***

9 F. Barbaresco, ***Koszul Information Geometry and Souriau Lie Group Thermodynamics***

11:15 – 11:30 Coffee break

11:30 – 12:45 Poster session 1

Chaired by: Kevin Knuth / Nicolas Gac / Mathiew Kowalski / Aurélia Fraisse

105 John Skilling, Cluster Analysis

28 Anthony Garrett, In search of hidden variables - Poster Presentation

25 Swapnesh Panigrahi, Julien Fade and Mehdi Alouini, Contrast enhancement in polarimetric imaging with correlated noise fluctuations

33 Xiao Yu , TIB parametrization of signal processing

34 Julianna Pinele, Joao Strapasson and Sueli Costa, An Upper Bound for the Fisher-Rao Distance of Multivariate Normal Models

36 Marcel Reginatto, Francis Gagnon-Moisan, Jorge Guerrero, Ralf Nolte and Andreas Zimbal, A Bayesian method to estimate the neutron response matrix of a single crystal CVD diamond detector

41 Francois Bertholon, Olivier Harant, Christian Jutten, Pierre Grangeat, Bertrand Bourlon and Laurent Gerfault, Signal analysis of NEMS sensors at the output of a chromatography column

59 Diego González and Sergio Davis, Inference of trajectories over a time-dependent phase space distribution

79 A. Heitor Reis, Entropy production between extremes, or how to reach equilibrium by the fastest way

50 Osamu Komori and Shinto Eguchi, Maximum power entropy method for ecological data analysis

86 Selman Ipek and Ariel Caticha, Entropic Dynamics of Relativistic Quantum Fields: What is a Particle?

12:45 – 14:00 Lunch break

14:00 – 15:00 Oral session 2: Entropy Foundations

Chaired by: A. Mohammad-Djafari / J. Skilling

Invited talk: Prof. Mikhail Gromov ***On the structure of Entropy (96)***

15:00 – 15:30 Coffee break

15:30 – 17:30 Oral session 3: Information Geometry

Chaired by: A. Caticha / Steeve Zozor / F. Nielsen

29 Robert Wolak and Michel Nguiffo Boyom, Foliations in Information Geometry

20 Mitsuhiro Itoh and Hiroyasu Satoh, Fisher Information Geometry of the Barycenter Map

37 Kei Kobayashi, Mitsuru Orita and Henry P. Wynn, Statistical analysis via the curvature of data spaces

75 Gregory Chirikjian and Bernard Shiffman, Entropy and Integral Geometry on Motion Spaces

18:00 – 19:30 Cocktail

Tuesday, September 23rd

09:00 – 10:45 Oral session 4 – History of Science

Chaired by: John Skilling / Antony Garrette

Invited talk: Stephano Bordoni, Duhem's abstract thermodynamics (89)

30 Olivier Rioul and José Carlos Magossi, Shannon's Formula and Hartley's Rule: A Mathematical Coincidence?

52 Anthony Garrett, From samples, how many classes are there in a population?

10:45 – 11:00 Coffee break

11:00 – 12:30 Oral session 5 – Bayesian inference

Chaired by: R. Bontekoe / F. Nielsen / J. Skilling

10 Fred Daum, Bayesian particle flow for estimation, decisions and transport

55 Jaehyung Choi and Andrew P. Mullhaupt Application of Kahler manifold to signal processing and Bayesian inference

87 Michael Habeck, Nested sampling with demons

12:30 – 14:00 Lunch break

14:00 – 15:30 Oral session 6 - Foundations and Geometry

Chaired by: John Skilling / F. Nielsen / Ariel Caticha

95 Jun Zhang, Information geometry and optimal transport theory

43 Marios Valavanides and Tryfon Daras Optimum Operating Conditions for Two-Phase Flow in Pore Network Systems: Conceptual /Numerical Justification Based on the MEP principle

8 Julio Stern, Cognitive-Constructivism, Quine, Dogmas of Empiricism, and Muenchhausen's Trilemma

15:30 – 18:30 Poster session 2

Chaired by: Jean-François Bercher / Valérie Girardin / Marcel Reginatto

4 Bontekoe (Demo avec Mathematica)

6 Hellinton Takada and Julio Stern, Non-Negative Matrix Factorization and Term Structure of Interest Rates

7 Hellinton Takada and Julio Stern, Information Criterion for Selection of Ubiquitous Factors

15 Alexis Decurninge and Frédéric Barbaresco, Robust Burg Estimation of stationary autoregressive mixtures covariance

17 Keiko Uohashi, Harmonic maps relative to α-connections on statistical manifolds

18 Soumia Sid Ahmed, Zoubeida Messali, Abdeldjalil Ouhabi, Sylvain Trepout, Cedric Messaoudi and Segio Marco Non Parametric Denoising Methods Based on Wavelets : Application to Electron Microscopy Images in Low Time Exposure

22 Youssef Bennani, Luc Pronzato and Maria-Joao Rendas, Most Likely Maximum Entropy for Population Analysis: a case study in decompression sickness prevention

46 Olivier Schwander, José Picheral, Nicolas Gac, Daniel Blacodon and Ali Mohammad-Djafari, Aero-accoustics source separation with sparsity inducing priors in the frequency domain

56 Ramandeep Johal, Renuka Rai and Guenter Mahler, Bounds on Thermal Efficiency from Inference

60 Yasmín Navarrete, Sergio Davis and Gonzalo Gutiérrez, Maximum entropy modelling of opinions in social groups

62 Mina Aminghafari and Maryam Hematti, Comparing Entropy and Energy Goodness-of-fit test for Lévy Distribution

63 Leila Gharsalli, Hacheme Ayasso, Bernard Duchêne and Ali Mohammad-Djafari, Variational Bayesian Approach with a heavy-tailed prior distribution for solving a non-linear inverse scattering problem

80 Abdelbassit Boualem, Meryem Jabloun, Philippe Ravier, Marie Naiim and Alain Jalocha, Assessment of two MCMC algorithms convergence for Bayesian estimation of the particle size distribution from multiangle dynamic light scattering measurements

81 Isaac Almasi and Adel Mohammadpour, Bayesian Reconstruction in Lévy Distribution

65 Ben Placek and Kevin Knuth, A Bayesian Analysis Of Kepler-2b Using The EXONEST Algorithm

77 Mohsen Salehi, Adel Mohammadpour and Mina Aminghafari,TFBS Prediction with Stochastic Differential Equation and Time Series

83 Mahdi Teimouri, Saeid Rezakhah and Adel Mohammadpour, Entropy-based goodness-of-fit test for positive stable distribution

88 Ali Ghaderi, On coarse graining of information and its application to pattern recognition

90 Veroni Jayawardana, Adom Giffin and Sumona Mondal,Bayesian analysis of factors associated with Fibromyalga syndrome subjects

91 Adom Giffin, Joseph Skufca and Palmer Lao, Using Bayes factors for multi-factor, biometric authentification

93 Steven H. Waldrip, Robert K. Niven, Markus Abel and Michael Schlegel MaxEnt Analysis of a Water Distribution Network in Canberra, ACT, Australia

Wednesday, September 24th

Oral session 7 (9h00-10h15) - Quantum physics

Chaired by: S. Zozor / J.F. Bercher / F. Barbaresco

Invited talk: **R. Balian, The entropy-based quantum metric (98)**

27 Ryszard Kostecki A new maximum entropy/information geometry approach to bayesian foundations of quantum theory

10:15 – 10:30 Coffee break

10:30 – 12:45 Oral session 8 – Quantum physics

Chaired by: Roger Balian / J.F. Bercher / Ariel Caticha / R. Kostecki

84 Ariel Caticha, Daniel Bartolomeo and Marcel Reginatto Entropic Dynamics: from Entropy and Information Geometry

38 Marcel Reginatto The geometrical structure of quantum theory as a natural generalization of information geometry

44 Stephan Weis The MaxEnt extension of a quantum Gibbs family, convex geometry and geodesics

45 Steeve Zozor, Gustavo Martín Bosyk, Mariela Portesi, Tristan Osán and Pedro Walter Lamberti Beyond Landau--Pollak and entropic inequalities: geometric bounds imposed on uncertainties sum

61 Klil Neori and Philip Goyal Anyons in the Operational Formalism

12:45 – 14:00 Lunch break

14:00 – 15:00 Oral session 9 - Quantum entropy

Chaired by: S. Zozor / R. Balian / Roman Belavkin

64 Roman Belavkin On Variational Definition of Quantum Entropy

23 Pierre Maréchal Duality for maximum entropy diffusion MRI

Social activities: see page XXXX

Visit of le Clos Lucé and Chenonceau castle

Gala Diner at Chissay Castle

Thursday, September 25th

9:00 – 10:30 Oral session 10
Geometric Structure of Information

Chaired by: Ariel Caticha / John Skilling

Invited talk: D. Bennequin, Topological forms of information (97)

24 Ke Sun and Stéphane Marchand-Maillet Information Geometry for Semi-parametric and Supervised Density Estimation

10:30 – 10:45 Coffee break

10:45 – 12:45 Oral session 11 Learning, Information, Divergence

Chaired by: Frank Nielson / Kevin Knuth / Fred Daum / O. Rioul

26 Jean-François Bercher, Valérie Girardin, Justine Lequesne and Philippe Regnault Goodness-of-fit tests based on (h,φ)-divergences and entropy differences

35 Takashi Takenouchi, Osamu Komori and Shinto Eguchi A novel boosting algorithm for multi-task learning based on the Itakuda-Saito divergence

12 Alexander Zuevsky Towards automorphic to differential correspondence for vertex algebras

82 Michel Broniatowski, Jana Jureckova and Amor Keziou Minimum risk equivariant minimum divergence estimates for moment condition models

12:45 – 14:00 Lunch break

14:00 – 15:30 Oral session 12 – Bayesian learning

Chaired by: Stephan Weis / F. Nielsen / Ariel Caticha

78 Frank Nielsen On learning statistical mixtures maximizing the complete likelihood

93 U. von Toussaint, Robust phase estimation for signals with a low signal-to-noise-ratio

73 Philippe Cuvillier Time-coherency of Bayesian priors of transient semi-Markov chains for sequential alignment

15:30 – 15:45 Coffee break

15:45 – 18:30 Poster session 3

Chaired by: Mathieu Kowalski / *Aurélia Fraisse* / Olivier Rioul / Sueli Costa

5 Justine Lequesne and Valérie Girardin, Analysis of information into marginal effects

13 Marco Congedo and Alexandre Barachant, A Special Form of SPD Matrix for Interpretation and Visualization of Data Manipulated with Riemannian Geometry

19 Jean-François Degurse, Information geometry for radar detection in heterogeneous environments

21 Alice Le Brigant, Marc Arnaudon and Frédéric Barbaresco, Probability on spaces of curves and in the associated metric spaces via information geometry: Application to the statistical study of non-stationary Radar signatures

31 Muhammed Sutcu and Ali Abbas, First Order Dependence Trees with Cumulative Residual Entropy

47 Steeve Zozor and Jean-Marc Brossier, De Bruijn identity: from Shannon entropy and Fisher information to generalized $f-$divergences and $f-$Fisher divergences

48 Anass Bellachehab and Jeremie Jakubowicz , Distributed consensus for metamorphic systems using a gossip algorithm for CAT(0) metric spaces

51 James Walsh and Kevin Knuth, Information-Based Physics, Influence and Forces

53 Mircea Dumitru and Ali Mohammad-Djafari, Estimating the periodic components of a biomedical signal through Inverse Problem modeling and Bayesian Inference with sparsity enforcing prior

54 Li Wang, Ali Mohammad-Djafari and Nicolas Gac, Bayesian 3D X-ray Computed Tomography image reconstruction with a Scaled Gaussian Mixture prior model

57 Geert Verdoolaege, Geodesic least squares regression for scaling studies in magnetic confinement fusion

66 Ning Chu, Ali Mohammad-Djafari, Nicolas Gac and José Picheral, A Hierarchical Variational Bayesian Approximation Approach in Robust Acoustic Imaging

68 Yannis Kalaidzidis, Inna Kalaidzidis and Marino Zerial, Quantitative microcopy: markers colocalization on intracellular vesicular structures

70 Kiamars Vafayi, Maximum-entropy method of moments for density estimation with large number of moments

72 Kevin Knuth, The Problem of Motion: The Statistical Mechanics of Zitterbewegung

74 Sergio Davis, Joaquín Peralta, Yasmín Navarrete, Diego González and Gonzalo Gutiérrez, A Bayesian interpretation of first-order phase transitions

94 R. Preuss and U. von Toussaint, Bayesian uncertainty quantification for an electrostatic plasma model

Friday, September 25th

09:00 – 11:00 Oral session 13 Entropy and Information Geometry

Chaired by: John Skilling/ F. Nielsen / F. Barbaresco

39 Subrahamanian K S Moosath and Harsha K.V., F-Geometry and the Geometry Induced by a Two Point Function on a Statistical Manifold

49 Luigi Malagò and Giovanni Pistone, Gradient flow of the stochastic relaxation on a generic exponential family

11 Harsha K V and Subrahamanian Moosath K S, Geometry of F-likelihood Estimators and F-Max-Ent Theorem

92 Robert Niven, Markus Abel, Steven H. Waldrip, Michael Schlegel and Bernd R. Noack MaxEnt Analysis of Flow and Reaction Networks

11:00 – 11:15 Coffee break

11:15 – 12:45 Oral session 14 - Bayesian inference

Chaired by: P. Maréchal / Steeve Zozor / J.F. Bercher

14 Bahruz Gadjiev and Tatiana Progulova Origin of generalized entropies and generalized statistical mechanics for superstatistical multifractal systems

40 Hiroshi Matsuzoe Information geometry of Bayesian statistics

103 R. Preuss and Udo V. Toussaint Bayesian uncertainty quantification for an electrostatic plasma model

12:45 – 14:00 Lunch break

14:00 – 17:00 Oral session 15 - Maximum Entropy Principle

Chaired by: John Skilling / Hiroshi Matsuzoe

3 Alexander Fradkov and Dmitry Shalymov Dynamics of non-stationary processes that follow the MaxEnt principle for Shannon and Renyi entropies

16 Patrick Bogaert and Sarah Gengler MinNorm approximation of MaxEnt/MinDiv problems for probability tables

42 Shinto Eguchi, Osamu Komori and Atsumi Ohara Duality in a maximum generalized entropy model

58 Diego González, Sergio Davis and Gonzalo Gutiérrez Newtonian Dynamics from the principle of Maximum Caliber

17:00 – 17:15 Coffee break

17:15 – 17:30 Closing Session - *A. Mohammad-Djafari & F. Barbaresco*

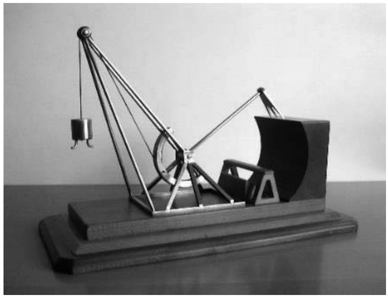
SOCIAL EVENTS

Guided Visit of Le Clos Lucé

16:00 – 18:00

Meeting point : at the registration desk

2 Rue du Clos Lucé, 37400 Amboise
Tel. : 00 33 (0)2 47 57 00 73



Leonardo da Vinci at Le Clos Lucé

Painter, inventor, engineer, scientist, humanist, philosopher, Leonardo da Vinci is the very incarnation of the universalist spirit of the Renaissance, and remains its main symbol. His life was punctuated by frequent periods staying at the Italian courts, until Francis I invited him to come and settle at the Château du Clos Lucé.

Leonardo da Vinci at Le Clos Lucé (1516-1519)

In 1515, Francis I won the Milan area in the battle of Marignan, and invited Leonardo to come to France. The following year, the artist arrived at the Château du Clos Lucé after a journey across the Alps on mule-back. He brought with him three paintings: the Mona Lisa, Saint John the Baptist, and The Virgin and Child with Saint Anne. Leonardo was named as “The King’s First Painter, Engineer and Architect”, and received a pension of 1000 gold crowns per year. In 1519, the artist died at Le Clos Lucé at the age of 67.

Guided Visit of Chenonceau Castle

18:15 Meeting point: at the registration desk and departure by bus

Audio guided visit from 18:30 to 19:30

Departure by bus at 20:00 for the Gala Dinner

Property of the Crown, then royal residence, Chenonceau Castle is an exceptional site not only because of its original design, the richness of its collections, its furniture and its decorations, but also because of its destiny, since it was loved, administrated and protected by women, who were all extraordinary and who, for the most part have marked history.

For the historical background, the “Château des Dames” was built in 1513 by Katherine Briçonnet, and successively embellished by Diane de Poitiers then Catherine de Medici. Chenonceau was protected from the hardship of the revolution by Madame Dupin.

Chenonceau Castle has an exceptional museum collection of the Old Masters’ paintings: Murillo, Le Tintoret, Nicolas Poussin, Le Corrège, Rubens, Le Primatice, Van Loo... as well as an extremely rare selection of Flanders Tapestries from the 16th century.

Throughout its history, this emblematic Castle has always attracted talent and inspired great artists. Conveying beauty and combining the elegance of architecture with that of the spirit is also sharing an elegant way of life.

iPhone and iPod touch applications.

The château of Chenonceau offers a free application, "Discover Chenonceau", which will give you all the necessary information for your next visit. Now available on App Store. The full visit of the château is available in eleven languages on App Store.

Gala Dinner at Chissay Castle

20:30

1 à 3 Place Paul Boncour, 41400 Chissay-en-Touraine

Tel. : 00 33 (0)2 54 32 32 01

Chissay Castle is a poem in stone brought to life.

This former royal residence in the shadow of Chenonceau castle is the ideal place to discover the wonderful Loire Valley.

You will enjoy the renaissance courtyard and your dinner served in the magnificent recently restored guards' room.

ABSTRACTS

Sunday, September 21st

09:00 - 10:00 Registration

10:00 – 13:00 Tutorial session 1

Chaired by Ali Mohammad-Djafari / Frédéric Barbaresco / John Skilling

1. Modern Probability Theory (71) - Kevin H. Knuth

A theory of logical inference should be all-encompassing, applying to any subject about which inferences are to be made. This includes problems ranging from the early applications of games of chance, to modern applications involving astronomy, biology, chemistry, geology, jurisprudence, physics, signal processing, sociology, and even quantum mechanics. This paper focuses on how the theory of inference has evolved in recent history: expanding in scope, solidifying its foundations, deepening its insights, and growing in calculational power.

2. The Basics of Information Geometry (85) - Ariel Caticha

A main concern of any theory of inference is to pick a probability distribution from a set of candidates and this immediately raises many questions. What if we had picked a neighboring distribution? What difference would it make? What makes two distributions similar? To what extent can we distinguish one distribution from another? Are there quantitative measures of distinguishability? The goal of this tutorial is to address such questions by introducing methods of geometry. More specifically the goal will be to introduce a notion of "distance" between two probability distributions.

A parametric family of probability distributions forms a statistical manifold, namely, a space in which each point represents a probability distribution. Generic manifolds do not come with a pre-installed notion of distance; such additional structure has to be purchased separately in the form of a metric tensor. Statistical manifolds are, however, an exception: a theorem due to N. Čencov (1981) states that up to an overall scale factor there is only one metric that takes into account the fact that these are not distances between simple structureless dots but distances between probability distributions.

To educate our intuition I will briefly sketch a couple of derivations of the information metric and provide a couple of examples. I will not develop the subject in all its possibilities but I will emphasize one specific result. Having a notion of distance means we have a notion of volume and this in turn implies that there is a unique and objective notion of a distribution that is uniform over the space of parameters—equal volumes are assigned equal probabilities. Whether such uniform distributions are maximally non-informative, or whether they define ignorance, or whether they reflect the actual *prior beliefs of any rational agent, are all important issues but they are quite beside the specific point to be made here: that these distributions are uniform and this is not a matter of subjective judgment but of objective mathematical proof.*

3. Voronoi diagrams in information geometry (94) - Franck Nielsen

Divergence functions, which are traditionally viewed as a bi-variate function on some manifold M, are here viewed as functions on the cross-manifold MxM which generate a statistical structure (a Riemannian metric plus a pair of torsion-free conjugate connections) along its diagonal manifold. Imposing compatibility conditions allow us to define a divergence function that is "proper". Further conditions can be imposed so that the cross-manifold may admit a complex representation, linking the divergence function to the "potential" on MxM. For the family of divergence functions induced by a convex function (Zhang, 2004), it is shown that they are proper and their Kahler potentials are given exactly by the inducing convex function. These results highlight the "reference-representational biduality" in Information Geometry.

14:30 – 19:30 Tutorial session 2

Chaired by Ariel Caticha / Ken H. Knuth / Frank Nielsen

4. Foundations and Geometry (104) - John Skilling

Probability theory has a solid foundation based on elementary symmetries, nowadays refined to associativity augmented with either commutativity or order. Few workers now contest the sum and product rules of standard probability calculus (often called "Bayesian" although there's no rational alternative and a unique form needs no adjective).
The practice of inference is understood and agreed.

Yet pure inference is not the end of the story. We may wish to simplify a distribution, or aggregate several into a single representative, in a minimally damaging way. For such purposes, we wish to know how far one probability distribution is from another, so that we can define what we mean by minimal damage.

Over the years, many candidate distances have been proposed, used, and generalised

to cover measures (distributions that can have any total) and even to positive matrices. On these distances, geometries have been constructed. Their very multiplicity, though, to say nothing of their ad hoc production, indicates that none has been found wholly satisfactory. There is a reason for that.

The reason is that there is only one connection that allows data from arbitrary

partitions of the coordinate space to be combined consistently. That connection is the unique information

H(p; q) = pi log pi/qi

known to statisticians as the Kullback-Leibler. And H is asymmetric,

H(p; q) not= H(q; p).

The asymmetry is both central and obvious. It can't be evaded. To pass from distribution q = (1/2 ; 1/2) to p = (1; 0) takes one bit of information (which might tell us that a coin was "heads"). But the reverse passage from (1; 0) representing a coin known to be "heads" to (1/2 ; 1/2) is impossible because "tails" is supposedly known to be false.

It follows that probability distributions do not form a metric space. Consequently,

all geometries must fail. More precisely, any proposal based on a symmetric distance l(p; q) = l(q; p) must be

opposed to the uniqueness of H and will fall foul of elementary criteria, thereby being open to counter-example. Specifically, the popular claim that the Fisher metric g = \nabla\nabla H defines generally applicable geodesic paths and lengths, with associated density

\sqrt{p |g|}, is to be resisted because it contradicts the foundation of H itself.

150 years ago, Riemann ditched Euclid. Perhaps it is time to ditch Riemann.

5. Uncertainty quantification for computer model (101) - Udo V. Toussaint

The quantification of uncertainty for complex simulations is of increasing importance as well as a significant challenge. Bayesian and non-Bayesian probabilistic uncertainty quantification methods like polynomial chaos (PC) expansion methods or Gaussian processes have found increasing use over the recent years. This contribution describes the use of Gaussian processes and collocation methods for the propagation of uncertainty in computational models using illustrative examples as well as real-world problems. In addition the existing challenges like phase-transitions are outlined.

6. Geometric Structures Induced From Divergence Functions (95) - Jun Zhang

Divergence functions are generalizations of cross-entropy; they characterize (non-symmetric) proximity of pairs of points of a vector space or smooth manifold in general. As surrogate to the symmetric metric function, divergence functions play important roles in statistical inference, machine learning, image processing, optimization, etc. This talk will review the various geometric structures induced from a divergence function defined on a manifold. Most importantly, a Riemannian metric with a pair of torsion-free affine connections can be induced on the manifold; this is the so-called “statistical structure” in Information Geometry. Additional structures may emerge depending on the functional form of the divergence. A general family of divergence functions can be constructed based on a smooth and strictly convex function, which unifies the various known families. Such divergence functions results in a manifold equipped with a pair of bi-orthogonal coordinates, and therefore Hessian structure, reflecting “reference-representation biduality”, and an equiaffine structure such that parallel volume forms exist. Computational advantages of this convex-based divergence functions will be discussed.

7. Koszul Information Geometry and Souriau Lie Group Thermodynamics (9)

Frédéric Barbaresco

The Koszul-Vinberg Characteristic Function (KVCF) is a dense knot in important mathematical fields such as Hessian Geometry, Kählerian Geometry, Affine Differential Geometry. This paper develops KVCF as the foundation of Information Geometry, transverse concept in Thermodynamics, in Statistical Physics and in Probability. From general KVCF definition, the paper defines Koszul Entropy, that coincides with the Legendre transform of minus the logarithm of KVCF (their gradients defining mutually inverse diffeomorphisms). These dual functions are compared by analogy in thermodynamic with dual Massieu-Duhem potentials. Hessian of minus the KVCF logarithm provides a non-arbitrary Riemannian metric for Information Geometry. We will observe the fundamental property that barycenter of Koszul Entropy is equal to Koszul entropy of barycenter. We present then a generalization of the characteristic function by physicist Jean-Marie Souriau in statistical physics, introducing the concept of co-adjoint action of a group on its momentum space, defining physical observables like energy, heat and momentum as pure geometrical objects. We will compare moment map with the dual coordinate in Koszul model (barycenter where entropy is maximum) and give a vector valued definition of Maximum Entropy. In covariant Souriau model, Gibbs equilibriums states are indexed by a geometric parameter, the Geometric Temperature, with values in the Lie algebra of the dynamical group, interpreted as a space-time vector (a vector valued temperature of Planck), giving to the metric tensor a null Lie derivative. Information Fisher metric appears as the opposite of the derivative of Moment map by geometric temperature, equivalent to a Geometric Capacity. We will synthetize the analogies between Koszul and Souriau models where Information Geometry is considered as a particular case of Koszul Hessian Geometry. Information Geometry metric is then characterized by invariances, by automorphisms of the convex cone or by Dynamic groups. We conclude, interpreting Legendre transform as Fourier transform in (Min,+) algebra, with new definition of Entropy given by the following relation: Entropy = - Fourier(Min,+) ° Log ° Laplace(+,X) .

8. Bayesian and Information Geometry in signal processing (1)

Ali Mohammad-Djafari

The main object of this tutorial article is first to review the main inference tools using Bayesian approach, Entropy, Information theory and their corresponding geometries.

This review is focused mainly on the ways these tools have been used in signal and

image processing.

After a short introduction of different quantities related to the Bayes rule, the entropy and the maximum entropy principle, we will study their use in different

fields of data and signal processing such as: source separation, model order

selection, spectral estimation and, finally, general linear inverse problems.

Monday, September 22nd

08:00 - 09:00 Registration

09:00 - 09:15 Opening Session

A. Mohammad-Djafari & F. Barbaresco

09:15 - 11:15 Oral session 1: Statistical Manifolds

Chaired by: Steeve Zozor / J.F. Bercher / Ariel Caticha /

67 Jean-Daniel Boissonnat, David Cohen-Steiner, Pooran Memari and Arnaud Poinas, ***Triangulating Statistical Manifolds***

In this paper, we describe an algorithm to construct an intrinsic Delaunay triangulation of a smooth closed statistical manifold, i.e. a manifold of probability density functions (PDFs). In differ- ent applications (namely in data analysis, information geometry, biology, shape analysis, compression, scientific visualization), statistical data are often seen as points in a metric space, which is most fre- quently high dimensional. The generally accepted hypothesis is that, although it is embedded in spaces of high dimensions, data lives usually close to a much smaller structure with a low intrinsic dimension. Making use of this intrinsic geometric structure, this paper will develop a certified algo- rithm that can reconstruct statistical manifolds with a complexity that depends only linearly on the ambient dimension.

We consider the Fisher information metric defined as a particular Riemannian metric on smooth statistical manifolds to calculate the informational difference between measurements. Interestingly, this metric can be understood as an infinitesimal form of the relative entropy of distributions; more specifically, it is the Hessian of the Kullback-Leibler divergence which is a particular instance of well-studied Bregman divergences for which efficient reconstruction algorithms have been proposed [1].

Our algorithm builds over existing algorithms for Bregman Delaunay triangulations [1] and mani- fold reconstruction using the tangential Delaunay complex proposed in [2]. The central idea is to define Bregman Delaunay triangulations locally and to glue these local triangulations together by removing inconsistencies between them. We view the inconsistencies as arising from instability in the Bregman Delaunay triangulations, and exploit and adapt the results presented in [3] for the Euclidean case. In particular, our technique heavily uses duality. Our main result is an algorithm whose complexity depends only linearly on the ambient dimension, and produces a Bregman Delaunay complex which is guaranteed to be a triangulation of the manifold under appropriate sampling conditions. We also demonstrate that the resulting triangulation coincides with the intrinsic Delaunay complex for the Fisher metric.

69 Nina Miolane and Xavier Pennec, ***Statistics on Lie groups : a need to go beyond the pseudo-Riemannian framework***

Lie groups appear in many fields from Medical Imaging to Robotics. In Medical Imaging and particularly

in Computational Anatomy, one often models an organ’s shape by the deformation of a reference shape, in

other words: by an element of a Lie group. Thus in order to model the variability of an organ’s shape, one

needs to perform statistics on Lie groups. A Lie group G is a manifold with an additional group structure. Statistics on manifolds have been studied throughout the literature, with the pioneer work of Frechet, Karcher and Kendall followed by many others. For the special case of Lie groups, it seems natural to require that the statistical notions are consistent with the group structure. Thus, one may wonder if we could use the results on pseudo-Riemannian manifolds to perform consistent statistics on Lie groups. In particular, one may ask if some pseudo-Riemannian metric could be used to characterize the bi-invariant mean on G introduced with the Cartan-Schouten connection. Answering this question implies constructing a bi-invariant pseudo-metric compatible with the group structure. This is precisely the purpose of our work. We show an algorithmic procedure that constructs bi-invariant pseudo-metrics on connected finite dimensional Lie groups. Our procedure relies on the classification theorem of Medina and Revoy on Lie groups that admit bi-invariant pseudo-metrics. We examine the indecomposable pieces of the adjoint representation of the corresponding Lie algebra g and construct bi-invariant pseudo-metrics on them, whether they are simple Lie algebras or double extensions of a smaller Lie algebra. The first case of simple Lie algebras is trivial because the space of bi-invariant pseudo-metrics is spanned by the Killing form. We treat the second case recursively. Our procedure also answers if the Lie group G given as input admits a bi-invariant pseudo-metric. Our experiments show that SE(3) admits bi-invariant pseudo-metrics but most Lie groups do not. For example the Heisenberg group, the group of scalings and translations, the group of unitriangular matrices do not admit any bi-invariant pseudo-metric while they do possess a bi-invariant mean. It proves that the pseudo-Riemannian framework is not rich enough to perform statistics on Lie groups. There is a need to investigate other geometric structures on Lie groups.

76 Max H. M. Costa, Chandra Nair and Olivier Rioul, ***From Almost Gaussian to Gaussian***

We consider lower and upper bounds of the difference of the differential entropies of a Gaussian random vector and an approximately Gaussian random vector after they are "smoothed" by an arbitrarily distributed random vector of finite power. These bounds are important to establish the optimality of the corner points in the capacity region of Gaussian interference channels. A problematic issue in a previous attempt to establish these bounds was detected in 2004 and the mentioned corner points have since been dubbed "the missing corner points". The importance of the given bounds comes from the fact that they induce Fano-type inequalities for the Gaussian interference channel. Usual Fano inequalities are based on a communication requirement. In this case, the new inequalities are derived from a non-disturbance constraint.

The upper bound on the difference of differential entropies is established by the Data Processing Inequality (DPI). For the lower bound, we argue that if follows from the DPI and a continuity argument.

9 F. Barbaresco, ***Koszul Information Geometry and Souriau Lie Group Thermodynamics***

The Koszul-Vinberg Characteristic Function (KVCF) is a dense knot in important mathematical fields such as Hessian Geometry, Kählerian Geometry, Affine Differential Geometry. This paper develops KVCF as the foundation of Information Geometry, transverse concept in Thermodynamics, in Statistical Physics and in Probability. From general KVCF definition, the paper defines Koszul Entropy, that coincides with the Legendre transform of minus the logarithm of KVCF (their gradients defining mutually inverse diffeomorphisms). These dual functions are compared by analogy in thermodynamic with dual Massieu-Duhem potentials. Hessian of minus the KVCF logarithm provides a non-arbitrary Riemannian metric for Information Geometry. We will observe the fundamental property that barycenter of Koszul Entropy is equal to Koszul entropy of barycenter. We present then a generalization of the characteristic function by physicist Jean-Marie Souriau in statistical physics, introducing the concept of co-adjoint action of a group on its momentum space, defining physical observables like energy, heat and momentum as pure geometrical objects. We will compare moment map with the dual coordinate in Koszul model (barycenter where entropy is maximum) and give a vector valued definition of Maximum Entropy. In covariant Souriau model, Gibbs equilibriums states are indexed by a geometric parameter, the Geometric Temperature, with values in the Lie algebra of the dynamical group, interpreted as a space-time vector (a vector valued temperature of Planck), giving to the metric tensor a null Lie derivative. Information Fisher metric appears as the opposite of the derivative of Moment map by geometric temperature, equivalent to a Geometric Capacity. We will synthetize the analogies between Koszul and Souriau models where Information Geometry is considered as a particular case of Koszul Hessian Geometry. Information Geometry metric is then characterized by invariances, by automorphisms of the convex cone or by Dynamic groups. We conclude, interpreting Legendre transform as Fourier transform in (Min,+) algebra, with new definition of Entropy given by the following relation: Entropy = - Fourier(Min,+) ° Log ° Laplace(+,X) .

11:30 – 12:45 Poster session 1

Chaired by: Kevin Knuth / Nicolas Gac / Mathiew Kowalski / Aurélia Fraisse

28 Anthony Garrett, In search of hidden variables –

Einstein was right: we should not be satisfied with the statistical character of the predictions of quantum mechanics when measuring an observable. It is the task of theoretical physicists to generate testable predictions. To be unable to do so is one thing, but to be satisfied with it is another. Physicists can predict the gyromagnetic ratio of the electron to 9 decimal places and analyse the first microseconds after the Big Bang, but cannot predict whether the next electron will go up or down in an inhomogeneous magnetic field in a Stern-Gerlach apparatus, even when every electron in the beam his the same wavefunction as a result of selection in a previous, differently-oriented Stern-Gerlach apparatus. The Stern-Gerlach experiment is now nearly a century old and it is time to look for hidden variables - not as a supplement to the wavefunction but as something transcending it. Every purported no-hidden-variable *theorem involves axioms, and - supposing it is correct technically - should be turned round and regarded as excluding hidden variables that conform to its axioms. This narrows the search. For example, thanks to Bell we know that the hidden variables must be nonlocal. Nonlocality was first posited by Isaac Newton in his theory of gravity, as action at a distance. The hidden variables must also be acausal, but even that is no cause for panic. Only seekers find (barring lucky accidents).*

25 Swapnesh Panigrahi, Julien Fade and Mehdi Alouini, Contrast enhancement in polarimetric imaging with correlated noise fluctuations

We quantify and analyse the gain in measurement precision in using polarimetric imaging as opposed to a simple intensity imaging in various situations. We demonstrate that when the background noise is sufficiently correlated, polarimetric imaging always out-performs a simple intensity imager. We have setup long range polarimetric imaging system over kilometric distance to study polarimetric imaging through fog. We intend to address the challenges in contrast enhancement, estimation and detection of a polarized beacon embedded in intense background using statistical and information theoretic techniques.

We describe a generalization of the TIB representation of a time invariant linear system from scalar (SISO) to vector (MIMO) case. Several advantageous features of the scalar case, such as numerical stability, fast state update and efficient dimensionality reduction, are also extended to the matrix case. Consider a d-dimensional innovations model,

$$\begin{aligned} z(t+1) &= Az(t) + Bx(t), \quad (1) \\ y(t) &= Cz(t) + x(t), \quad (2) \end{aligned}$$

with $y(t)$ being a sequence of d-dimensional measurement vectors and $z(t)$ being n-dimensional state vectors. By appropriate choice of state-space coordinates we choose A to be a lower triangular, and input balanced: $AA^* + BB^* = I$. As a special case of orthogonal filter, input balance ensures good numerical performance, and the triangular realization has high precision location of eigenvalues and a band matrix fraction representation which affords fast algorithms and a very efficient model reduction algorithm for both SISO and MIMO cases is also known for the TIB parametrization.

According to Douglas-Shaprio-Shields factorization, any strictly proper rational stable function can be factored into an unstable part and rational lossless part, where for the TIB realization the unstable part is related to C and the lossless part is associated with A and B. Once the lossless part has been determined, the transfer functions have convenient geometry. The Schur algorithm constructs scalar rational lossless functions, the extension to the MIMO case is the Schur tangential algorithm,

$$\mathcal{B}_i (1/\bar{w}_i)\, u_i = v_i, \quad (3)$$

where $\mathcal{B}_i$ is a sequence of rational lossless functions with degree/McMillan degree i. When $v_i = 0$, then w_i are the poles of the system, and $\mathcal{B}_i$ are the Blaschke-Potapov factorization in MIMO case (and Blaschke factors in the SISO case). The TIB pair (A, B) is uniquely determined by the poles of the system for SISO, whereas for MIMO, the null vectors u_i are also required to determine the lossless part. The reduction algorithm we propose is an hybrid one, which obtains the lossless part by minimizing the Hankel (H^∞) norm and the unstable part by minimizing H^2 norm. The lossless part (A, B) can be determined from a partial

SVD of the Hankel matrix given by

$$H_{ij} = f_{i+j-1}, i, j = 0, 1, \dots \tag{4}$$

where $\{f_i\}$ are the impulse responses. C can be therefore obtained by a well conditioned least squares regression.

In the SISO case, the choice of model parameters as the power series coefficients of the logarithm of the transfer function results in the Fisher information matrix is an identity matrix, providing a Euclidean statistical (Hilbert) manifold. Denoting $\log f(z) = a_0 + a_1 z + a_2 z^2 + \cdots$, since f and $\log f$ has the same singularity, we can apply our reduction algorithm on the alternative Hankel matrix

$$A_{ij} = a_{i+j-1}, i, j = 0, 1, \dots \tag{5}$$

and then recover the parameters after exponentiation. Mullhaupt and Choi proved in the SISO case, the information geometry is determined by the unstable part of the transfer function, which our algorithm can extend to the MIMO case.

References

[1] A.P.Mullhaupt, K.S.Riedel, *Fast Identification of Stable Innovation Filtersg*

[2] A.P.Mullhaupt, K.S.Riedel, *Band Matrix Representation of Triangular Input Balanced Form*

[3] A.P.Mullhaupt, K.S.Riedel, *Low Grade matrices and matrix fraction representations*

[4] A.P.Mullhaupt, K.S.Riedel, *Exponential Condition Number of Solutions Discrete Lyapunov Equation*

[5] B.Hanzon, M.Olivi, R.M.Peeters, *Balanced realizations of discrete-time stable all-pass systems and the tangential Schur algorithm*

[6] R.G.Douglas, H.S.Shapiro, A.L.Shields, *Cyclic vectors and invariant subspaces for the backward shift operator*

[7] J.Marmorat, M.Olivi, B.Hanzon, R.M.Peeters, *Matrix rational H^2 approximation: a state-space approach using Schur parameters*

34 Julianna Pinele, Joao Strapasson and Sueli Costa, An Upper Bound for the Fisher-Rao Distance of Multivariate Normal Models

Information geometry is approached here by considering the statistical model of multivariate normal distributions as a Riemannian manifold with the natural metric provided by the Fisher information matrix. Explicit forms for the Fisher-Rao distance associated to this metric and for the geodesics of general distribution models are usually very hard to determine. In the case of general multivariate normal distributions lower and upper bounds have been derived. We approach here some of these bounds and introduce a new one discussing their tightness in specific cases.

36 Marcel Reginatto, Francis Gagnon-Moisan, Jorge Guerrero, Ralf Nolte and Andreas Zimbal, A Bayesian method to estimate the neutron response matrix of a single crystal CVD diamond detector

Detectors made from artificial chemical vapor deposition (CVD) single crystal diamond have shown great potential for fast neutron spectrometry. The result of a measurement is a pulse height spectrum (PHS) which contains information about the energy spectrum of the incident neutrons. Deconvolution methods can be used to extract this information, but for this a response matrix is required. Current particle transport codes, while able to provide important information, are of limited use because they cannot simulate neutron responses of CVD diamond detectors that are of high enough quality to be used for deconvolution. Furthermore, the physics of the detector is not completely understood. Therefore, we propose instead a Bayesian method to estimate the response matrix which uses a combination of measurements (these were carried out at the PIAF accelerator facility of the Physikalisch-Technische Bundesanstalt) and numerical simulations.

The data consists of eight PHS of measurements of quasi-monoenergetic neutron beams with energies in the range of 7 MeV to 16 MeV. The aim of the analysis is a complete response matrix for this neutron energy range. We use the fact that the PHS can be modeled in terms of two additive components, a smooth "background" on which there is superimposed a "signal" with fine structure (e.g., peaks, shoulders). Both of these components are due to the neutrons incident on the detector, but they involve different types of interactions in the detector and thus have different characteristics. Our approach uses Bayesian signal-background separation techniques to carry out a joint estimate of the signal and background. Once the posteriors for the components have been calculated, the problem of estimating the response matrix for values of the neutron energy that lie in the gaps between measurements can be carried out for each component separately, which provides a great simplification. The interpolation of the background is relatively straightforward because it is a smooth function, which we model using radial basis functions. The interpolation of the fine structure that appears in the signal requires more care, and for this we use information from the numerical simulations. The method that we propose is quite general and it can be applied to other particle detectors with PHS that have similar characteristics.

41 Francois Bertholon, Olivier Harant, Christian Jutten, Pierre Grangeat, Bertrand Bourlon and Laurent Gerfault, Signal analysis of NEMS sensors at the output of a chromatography column

Gas chromatography is a technique to separate chemical components in gas state [1]. A chromatogram is a succession of peaks each corresponding to the output from the column of molecules of the same component. We consider new devices where nano-chromatography columns carved on silicium chip are coupled with sensors called NEMS, Nano Electro-Mechanical Systems. Those gravimetric sensors are vibrating cantilevers, covered with a chemical layer for molecular adsorption. The resonance frequency is controlled by a Phase Locked Loop control [2]. The output signal is the instantaneous frequency of the vibrating beam. This resonance frequency decreases when molecules are adsorbed on the cantilever. The frequency shift is proportional to the mass adsorbed. In this paper, we consider an inverse problem approach to retrieve the original gas mixture composition from the output signal based on Bayesian source separation [3] and model inversion.

The gas chromatography (GC) column and the NEMS sensors are here described in a molecular point of view through a stochastic model based on the random walk molecular model proposed by Giddings and Eyring [4,5]. Our proposed general model describes the shape of the signal output as the probability distribution of the retention time of the molecules within the column. Each acquisition is settled with some prior parameters defined by the known column and chemical properties such as the prior distributions for a molecule to stay within the mobile phase or to remain fixed on the stationary phase, or the prior distribution to be absorbed on the cantilever and then to be released in the mobile phase. Those parameters control the retention time distribution for each random walk of each molecule within the column. The inference on the unknown profile parameters is processed according to a Bayesian scheme [6,7] based on the joint probability density function of the output retention times.

Such inference requires first to identify each peak by locating their relative position. The position and width of each peak which define the parameters of our model are then estimated with a Maximum Likelihood Estimation to fully characterize the shape function of each peak. Then the mixture coefficients can be estimated. Indeed, each shape function corresponds to one gas and constitutes a base vector of elementary components. The scalars of this vector space which need to be determined are the proportions of each constituent inside the mixture. This defines a hierarchical model with 2 levels: component and GC signal. Then we propose a Hierarchical Bayesian source separation method to estimate the basis vector and the scalars of a given chemical sample in this basis. Finally we provide experimental evaluation on the analysis of a mixture of hydrocarbons.

59 Diego González and Sergio Davis, Inference of trajectories over a time-dependent phase space distribution

We present a numerical implementation of Maximum Caliber inference, based on the Metropolis-Hastings algorithm applied to trajectories. We compare the results of this *procedure with finite-differences simulation of the Fokker-Planck equation derived from a given Lagrangian Maximum Caliber problem (arXiv 1404.3249), in which the constraining function depends only on instantaneous position and velocity. We show that is it possible to recover the most probable trajectory and the values of the classical action from simulation of the Fokker-Planck equation for a given Hamiltonian.*

79 A. Heitor Reis, Entropy production between extremes, or how to reach equilibrium by the fastest way

In a recent paper the author has shown that the principles of extremum of entropy production are not first principles, instead they result from the maximization of conductivities under appropriate constraints. In this way, the maximum rate of entropy production (MEP) occurs when all the forces in the system are kept constant. On the other hand, the minimum rate of entropy production (mEP) occurs when all the currents that cross the system are kept constant.*

In this paper, we further develop the idea that entropy production rate is ruled by the more general principle of "maximization of flow access", which is mathematically translated to maximization of global system conductivities, and physically corresponds to the flow configuration that allows fluids to reach equilibrium the fastest.

While equilibrium states are defined through the maximum of entropy compliant with the existing constraints, transition between equilibrium states appears to be ruled by the minimum transition time compatible with the existing constraints. In this way, depending on the flow constraints, entropy production rate may assume any value in the range between the maximum and minimum values. We explore the consequences of this idea with regard both to the climate system and living systems.

(*) A. Heitor Reis, "Use and validity of principles of extremum of entropy production in the study of complex systems", Annals of Physics, 346, 22-27 (2014).

50 Osamu Komori and Shinto Eguchi, Maximum power entropy method for ecological data analysis

In ecology predictive models of the geographical distribution of certain species are widely used to capture the spatial diversity. Recently a method of Maxent based on Gibbs distribution is frequently employed to have reasonable accuracy of a target distribution of species at a site using environmental features such as temperature, precipitation, elevation and so on. It requires only presence data, which is a big advantage to the case where absence data is not available or unreliable. It also incorporates our limited knowledge into the model about the target distribution such that the expected values of environmental features are equal to the empirical average. Moreover, the visualization of the inhabiting probability of species is easily done with the aid of geographical coordination information from Global Biodiversity Inventory Facility (GBIF) in a statistical software R. However, the maximum entropy distribution in Maxent is derived from the Boltzmann-Gibbs-Shannon entropy, which causes unstable estimation of the parameters in the model when some outliers in the data are observed. To overcome the weak point and to have deep understandings of the relation *among the total number of species, the Boltzmann-Gibbs-Shannon entropy and Simpson's index, we propose a maximum power entropy method based on beta-divergence. It includes the Boltzmann-Gibbs-Shannon entropy as a special case, so it could have better performance of estimation of the target distribution by appropriately choosing the value of the power index beta. We demonstrate the performance of the proposed method by simulation studies as well as publicly available real data.*

86 Selman Ipek and Ariel Caticha, Entropic Dynamics of Relativistic Quantum Fields: What is a Particle?

Entropic Dynamics is an information-based framework that seeks to derive the laws of physics as an application of the methods of entropic inference. The dynamics is derived by maximizing an entropy subject to the appropriate constraints—the physically relevant information that the motion is continuous and non-dissipative.

In previous work we concentrated on the non-relativistic quantum theory which led to several new insights including the entropic nature of time; the interpretation of the phase of the wave function and of gauge transformations; the role of Hilbert spaces and complex numbers; the quantum measurement problem and the uncertainty relations.

In this paper we explore the relativistic regime and develop the quantum theory of scalar fields. With an appropriate choice of microstates and the use of concepts of information geometry we achieve a particularly direct derivation of Hamiltonian dynamics. For specific choices of the potential and a quantum potential motivated by information geometry we recover the functional Schroedinger equation for the Klein-Gordon field which is formally equivalent to the standard Fock space formalism. We study the configurations of lowest energy in order to gain insight into the nature of particles: What, after all, is a particle? What distinguishes those configurations that we call one-particle states from those configurations we call the vacuum? The related question of the nature of anti-particles is addressed by extending the analysis to complex scalar fields and to local U(1) gauge transformations.

A very appealing aspect of the entropic framework is that it leads to a new interpretation of the infinities that plague all formulations of quantum field theory. We find that what diverge are not the actual physical quantities but our uncertainties about them. These are not physical but epistemic effects; they are indications that the information that is relevant for certain predictions has not been incorporated into the analysis.

14:00 – 15:00 Oral session 2: Entropy Foundations

Chaired by: A. Mohammad-Djafari / J. Skilling

Invited talk : Prof. Mikhail Gromov ***On the structure of Entropy (96)***

Mathematics is about "interesting structures". What make a structure interesting is an abundance of interesting problems; we study a structure by solving these problems. The worlds of science, as well as of mathematics itself, is abundant with gems (germs?) of simple beautiful ideas. When and how may such an idea direct you toward *beautiful mathematics? I present in this talk a 20th century mathematician's perspective on Boltzmann's idea of entropy.*

15:30 – 17:30 Oral session 3: Information Geometry

Chaired by: A. Caticha / Steeve Zozor / F. Nielsen

29 Robert Wolak and Michel Nguiffo Boyom, Foliations in Information Geometry

Foliations appear very naturally in the spaces considered in Information Geometry. In the paper we propose to present an introduction to the theory of those foliations which can be encountered in Information Geometry. We will develop the theory of foliation, both regular and singular, on Hessian manifolds as well as the theory of foliations transversely modelled on Hessian manifolds.

Particular attention will be paid to foliation in small dimensions where the classification of those foliations is possible.

20 Mitsuhiro Itoh and Hiroyasu Satoh, Fisher Information Geometry of the Barycenter Map

We would like to report Fisher information geometry of the barycenter map associated with normalized Busemann function on an Hadamard manifold X and to present an application to Riemannian geometry of X from viewpoint of Fisher information geometry. This report is an improvement of our presentation at GSI13, Paris,2013 together with a fine investigation of the barycenter map relative to geodesics in the space of probabilities.

37 Kei Kobayashi, Mitsuru Orita and Henry P. Wynn, Statistical analysis via the curvature of data spaces

It has been known that the curvature of data spaces plays a role in data analysis. For example, the Fréchet mean (intrinsic mean) always exists uniquely for a probability measure on a non-positively curved metric space. In this talk, we use the curvature of data spaces in a novel manner. A methodology is developed for data analysis based on empirically constructed geodesic metric spaces. The population version defines distance by the amount of probability mass accumulated on traveling between two points and geodesic metric arises from the shortest path version. Such metrics are then transformed in a number of ways to produce families of geodesic metric spaces. Empirical versions of the geodesics allow computation of intrinsic means and associated measures of dispersion. A version of the empirical geodesic is introduced based on the graph of some initial complex such as the Delaunay complex. For certain parameter ranges the spaces become CAT(0) spaces and the intrinsic means are unique. In the graph case a minimal spanning tree obtained as a limiting case is CAT(0). In other cases the aggregate squared distance from a test point provides local minima which yield information about clusters. This is particularly relevant for metrics based on so-called metric cones which allow extensions to CAT(k) spaces. We show how our methods work by using some actual data.

Let X denote n-dimensional Euclidean space, let G denote the group of orientation-preserving rigid-body motions of X, and let Γ be a crystallographic (discrete co-compact) subgroup of G. In classical integral geometry, the Principal Kinematic Formula computes the integral [1, 3, 4]

$$\int_G \chi(A \cap gB)dg = \sum_{i=0}^{n} c_{in}\mu_i(A)\mu_{n-i}(B) \tag{1}$$

where $\chi(\cdot)$ denotes the Euler-Poincaré characteristic, dg is the Haar measure for G, c_{in} are known real constants, $\mu_i(\cdot)$ is the i^{th} invariant curvature measure of a body bounded by a smooth, closed, orientable surface, and A and B are arbitrary smooth bodies in X. Moreover, when the bodies are convex, $\chi(\cdot)$ can be replaced with the set indicator function, $i(\cdot)$. Under mild constraints on the bodies, an analogous formula holds for computing the volume in G corresponding to one convex body, B, moving inside another, C, without their bounding surfaces colliding [6]. Both formulas can be used together to compute the *change in entropy* of a freely moving convex body, B, moving inside a body, C, when an obstacle, A, is introduced in such a way that body B can circumvent it at any orientation without becoming jammed.

This is true because if V denotes the volume of free motion, then the probability of finding the moving body in a feasible region of configuration space is $1/V$, and the entropy is then $S = \log V$, and the change in entropy resulting from the presence of an obstacle is $\Delta S = \log V_1(B,C) - \log V_2(A,B,C)$.

We seek a formula analogous to (1) for the coset space $\Gamma\backslash G$ as a way to compute the entropy associated with all configurations of bodies arranged with crystallographic symmetry. This coset space is a smooth manifold of dimension $n(n+1)/2$, but because Γ is not normal in G, $\Gamma\backslash G$ is not a group. $\Gamma\backslash G$ is called a *motion space* [2]. Though it is not a group, a mapping $\pi : (\Gamma\backslash G) \times X \to X$ can be defined for each $\Gamma g \in \Gamma\backslash G$ and $x \in X$ as

$$\pi[\Gamma g, x] \doteq (\Gamma g)x = \bigcup_{\gamma\in\Gamma} \gamma \cdot (g \cdot x).$$

We derive a formula for the volume in $\Gamma\backslash G$ corresponding to arrangements such that $\pi[\Gamma g, B]$ is a collision-free configuration of bodies. That is, we derive a closed-form expression analogous to (1) for the integral

$$V_1 = \int_{F_{\Gamma\backslash G}} i\left((g \cdot B) \cap \bigcup_{\gamma \in \Gamma - e} (\gamma \circ g) \cdot B\right) dg$$

where $F_{\Gamma\backslash G}$ is a fundamental domain for the action of Γ on G, V_1 is the volume in $\Gamma\backslash G$ corresponding to bodies in collision, and $V_2 = V(F_{\Gamma\backslash G}) - V_1$ is the volume of the free space. $S(\Gamma, B) = \log V_2$ is the entropy of all possible collision-free crystallographic arrangements of copies of body B with symmetry group Γ. Moreover, we use the derived formula to examine changes in entropy as a function of the shape and size of the body B, and we point toward applications in the emerging field of "geometric frustration" [5].

References

[1] Chern, S.-S., "On the Kinematic Formula in the Euclidean Space of N Dimensions," *American Journal of Mathematics*, Vol. 74, No. 1 (Jan., 1952), pp. 227-236.

[2] Chirikjian, G.S., Yan, Y., "Mathematical Aspects of Molecular Replacement: II. Geometry of Motion Spaces," *Acta. Cryst. A* (2012).

[3] Klain, D.A., Rota, G.-C., *Introduction to Geometric Probability*, Cambridge University Press, 1997.

[4] Schneider, R., Weil, W., *Stochastic and Integral Geometry*, Springer-Verlag, Berlin, 2008.

[5] Sadoc, J.-F., Mosseri, R., *Geometrical Frustration*, Cambridge University Press, 2006.

[6] Yan, Y., Chirikjian, G.S., "Closed-form characterization of the Minkowski sum and difference of two ellipsoids," *Geometriae Dedicata* (to appear)

Tuesday, September 23rd

09:00 – 10:15 Oral Session 4 – History of Science

Chaired by: John Skilling / Antony Garrette

Invited talk: Stephano Bordoni, Duhem's abstract thermodynamics (89)

30 Olivier Rioul and José Carlos Magossi, Shannon's Formula and Hartley's Rule: A Mathematical Coincidence?

What has rapidly become the emblematic classical expression of information theory is Shannon's formula $C = 1/2 \log (1+ P/N)$ for the information capacity of a communication channel with signal-to-noise ratio P/N. Hartley's name is often associated to the formula, owing to Hartley's law: counting the highest possible number of distinguishable values for a given amplitude A and precision $\pm\Delta$ yields a similar expression $C'= \log (1+ A/\Delta)$.

In the information theoretic community, the following "historical" statements are generally well accepted: (1) Hartley did put forth his Hartley's rule twenty years before Shannon; (2) Shannon's formula as a fundamental tradeoff between transmission rate, bandwidth, and signal-to-noise ratio came out unexpected in 1948: times were not even ripe for this breakthrough; (3) Hartley's rule is inexact while Shannon's formula is characteristic of the additive white Gaussian noise (AWGN) channel; (4) Hartley's rule is an imprecise relation that is not an appropriate formula for the capacity of a communication channel. We show that all these four statements are questionable, if not wrong.

A detailed mathematical analysis explains the mathematical coincidence $C=C'$. We construct a sequence of additive noise channels, starting the uniform noise channel giving Hartley's rule, and converging towards the Gaussian channel, for which all channel capacities are given by the same Shannon's formula.

52 Anthony Garrett, From samples, how many classes are there in a population?

From samples of a population which may represent various classes (for example, nationalities), how many classes are actually represented in that population? The Bayesian solution to this question is presented. Specifically, the posterior probability distribution for the number of classes represented is given in terms of the prior probability distribution for the number of individuals in the population from every class. The posterior distribution for these numbers is calculated from Bayes' theorem using a multinomial likelihood to model the sampling process; from that posterior the number of classes represented can be extracted as the number of classes having nonzero representation. The result is given for sampling with and without replacement. This analysis supposes that all possible classes are specified. When that is not the case -- in particular when we do not know the number of possible classes -- then progress can still be made if we have information about the statistical properties of classes. We can then label the classes implicitly according to some ordering variable. Classes observed during sampling comprise an unknown permutation of that labelling, but this lack of knowledge can be marginalised over. The analysis can now be generalised to an unknown number of classes by applying a prior distribution for their number and marginalising over it.

11:00 – 12:30 Oral session 5 – Bayesian inference

Chaired by: R. Bontekoe / F. Nielsen / J. Skilling

10 Fred Daum, Bayesian particle flow for estimation, decisions and transport

We derive a new algorithm to compute Bayes' rule, using particle flow rather than a pointwise multiplication. The particle flow is induced by a log-homotopy of the conditional density from the prior to the posteriori. This algorithm is similar to Monge-Kantorovich transport, but it is much simpler and faster, because we do not use any optimality criteria. In particular, there is no relevant notion of physical effort to transport our particles. Consequently, we do not pick the flow to minimize a convex functional, and hence we avoid solving a variational problem. Moreover, we do not use resampling of particles or proposal densities or importance sampling or any Markov chain Monte Carlo method. But rather, we design the particle flow with the solution of a linear first order highly underdetermined PDE. We solve our PDE analytically as a formula, rather *than using numerical methods, in order to reduce computational complexity. We have many methods to solve this PDE, but the best methods all avoid explicitly computing the normalizing constant of the conditional density. Our theory applies to smooth nowhere vanishing probability densities, and it is similar to Jürgen Moser's coupling (AMS Transactions 1965). Our new particle flow exploits the exponential family of probability densities, which is a maximum entropy family and which has fixed finite dimensional sufficient statistics for nonlinear filters (Daum 1987). Nevertheless, our theory can solve arbitrary non-optimal transport problems, assuming smooth nowhere vanishing densities. We compare the new flow with state-of-the-art methods, including: Hamiltonian Monte Carlo, Metropolis-adjusted Langevin (MALA), regularized particle filters, auxiliary particle filters, etc. Our particle flow is many orders of magnitude faster than standard particle filters for the same accuracy. Furthermore, our filter beats the extended Kalman filter accuracy by several orders of magnitude for difficult nonlinear problems. Our algorithm avoids the well known problem of particle degeneracy in standard particle filters, and it reduces the computational complexity by many orders of magnitude for high dimensional fully-coupled problems (d = 2 to 42). The computational complexity of the flow for nonlinear filtering depends on: dimension of the state vector, stability of the plant, initial uncertainty of the state vector, measurement accuracy, shape of the conditional density (log-convex or multi-modal, etc.), Lipshitz constants of the log-densities, smoothness of the densities, and ill-conditioning of the Fisher information matrix.*

55 Jaehyung Choi and Andrew P. Mullhaupt Application of Kahler manifold to signal processing and Bayesian inference

We review the information geometry of linear systems and Bayesian inference, and the simplification available in the K\"ahler manifold case. We find conditions for information geometry of linear systems to be K\"ahler, and the relation of the K\"ahler potential to information geometric quantities such as $\alpha $-divergence, information distance and the dual $\alpha $-connection structure. The K\"ahler structure simplifies the calculation of the metric tensor, connection, Ricci tensor and scalar curvature, and the $\alpha $-generalization of the geometric objects. The Laplace-Beltrami operator is also simplified in the K\"ahler case. One of the goals in information geometry is the construction of Bayesian priors outperforming the Jeffreys prior, which we use to demonstrate the utility of the K\"ahler structure.

87 Michael Habeck, Nested sampling with demons

This article looks at Skilling's nested sampling from a physical perspective and interprets it as a microcanonical demon algorithm. Using key quantities of statistical physics we investigate the performance of nested sampling on complex systems such as Ising, Potts and protein models. We show that releasing multiple demons helps to smooth the truncated prior and eases sampling from it because the demons keep the particle off the constraint boundary. For continuous systems it is straightforward to extend this approach and formulate a phase space version of nested sampling that benefits from correlated explorations guided by Hamiltonian dynamics.

14:00 – 15:30 Oral session 6 - Foundations and Geometry

Chaired by: John Skilling / F. Nielsen / Ariel Caticha

95 Jun Zhang, Information geometry and optimal transport theory

Divergence functions are generalizations of cross-entropy; they characterize (non-symmetric) proximity of pairs of points of a vector space or smooth manifold in general. As surrogate to the symmetric metric function, divergence functions play important roles in statistical inference, machine learning, image processing, optimization, etc. This talk will review the various geometric structures induced from a divergence function defined on a manifold. Most importantly, a Riemannian metric with a pair of torsion-free affine connections can be induced on the manifold; this is the so-called "statistical structure" in Information Geometry. Additional structures may emerge depending on the functional form of the divergence. A general family of divergence functions can be constructed based on a smooth and strictly convex function, which unifies the various known families. Such divergence functions results in a manifold equipped with a pair of bi-orthogonal coordinates, and therefore Hessian structure, reflecting "reference-representation biduality", and an equiaffine structure such that parallel volume forms exist. Computational advantages of this convex-based divergence functions will be discussed.

43 Marios Valavanides and Tryfon Daras Optimum Operating Conditions for Two-Phase Flow in Pore Network Systems: Conceptual /Numerical Justification Based on the MEP principle

The mechanistic model DeProF considers steady-state two-phase flow in porous media as a composition of three flow patterns: connected-oil pathway flow, ganglion dynamics and drop traffic flow. The key difference is the degree of disconnection of the non-wetting phase which affects the relative magnitude of the rate of energy dissipation caused by capillary effects, compared to that caused by viscous stresses. An appropriate mesoscopic scale analysis leads to the determination of all the internal flow arrangements of the basic flow patterns that are compatible to the externally imposed flow conditions. The observed macroscopic flow is an average over the canonical ensemble of the flow arrangements.

Extensive DeProF simulations revealed that there exist a continuous line [a locus, r(Ca)] in the domain of the process operational variables -the capillary number, Ca, and the oil-water flowrate ratio, r- on which the efficiency of the process (oil produced per kW dissipated in pumps) attains local maxima. Such maxima have been experimentally identified.*

Subsequently, the existence of the locally optimum operating conditions could be rationally justified by the following conceptual inference. Steady state two-phase flow in porous media is an off-equilibrium process. The rate of global entropy production (a measure of the process spontaneity) is the sum of two components: the rate of mechanical energy dissipation at constant temperature (thermal entropy), Q/T, and a Boltzmann-type statistical-entropy production component, kDeProF lnΦ, directly related to the number of different physically admissible internal flow arrangements, Φ, associated with every flow condition (configurational entropy). By applying the MEP principle we may infer that optimum operation of the process is met on a locus of conditions whereby the process total entropy production rate takes maximum values.

To reduce the falsifiability of that inference, one needs to provide numerical evidence. To do so, it is necessary to deliver: (a) an efficient analytical/numerical scheme, to evaluate the number Φ of the different flow arrangements; (b) an expression for the constant kDeProF in the Boltzmann-entropy expression.

Combinatorial considerations provided the analytical background to evaluate the number of different micro-arrangements of the flow per physically admissible solution. The limiting procedure based on Stirling's approximation has been applied to downscale the excessively large computational effort associated with the numerical handling of operations between large factorial numbers. Still, an appropriate application of the Boltzmann principle needs to be implemented, to deliver an expression for the constant kDeProF pertaining to the sought process.

8 Julio Stern, Cognitive-Constructivism, Quine, Dogmas of Empiricism, and Muenchhausen's Trilemma

The Bayesian research group at University of Sao Paulo has been exploring a specific version of Cognitive Constructivism - Cog-Con - that has, among its most salient features, a distinctive "objective" character.

Cog-Con is supported by a specially designed measure of statistical significance, namely, ev(H | X) - the Bayesian epistemic value of sharp hypotheses H, given the observed data X.

This article explores possible parallels or contrasts between Cog-Con and the epistemological framework developed by the philosopher Willard van Orman Quine.

15:30 – 18:30 Poster session 2

Chaired by: Jean-François Bercher / Valérie Girardin / Marcel Reginatto

4 Bontekoe (Demo avec Mathematica)

Although the formalism of Bayesian analysis is elegant and compelling, the computation of the probability densities can be intricate. Only with the recent proliferation of computing power, Bayesian methods became doable in practice. A large number of specialised packages have been developed.

In this talk Bayesian analysis with Mathematica is discussed. Mathematica is unique in its integrated symbolic and numerical solving capacities. One consequence of this is that the code can be much shorter. Prof. Phil Gregory states "…, the time required to develop and test programs with Mathematica is approximately 20 times shorter than the time required to write and debug the same program in Fortran or C, …" .

However, Mathematica has a steep learning curve. It is a large language with thousands of functions from all corners of mathematics. Furthermore, the program texts themselves appear often terse. Learning Mathematica involves a considerable investment in time.

An essential part of the development of programs is the demonstration of its correctness. In this respect, the graphical capabilities of Mathematica are an excellent tool. In this talk the solution of a non-trivial Bayesian problem is demonstrated by *graphical means. A number of pitfalls are shown as well. The Mathematica code is provided as a take-off for new users.*

6 Hellinton Takada and Julio Stern, Non-Negative Matrix Factorization and Term Structure of Interest Rates

Non-Negative Matrix Factorization (NNMF) is a technique for dimensionality reduction with a wide variety of applications from text mining to identification of concentrations in chemistry. NNMF deals with non-negative data and results in non-negative factors and factor loadings. Consequently, it is a natural choice when studying the term structure of interest rates. In this paper, NNMF is applied to obtain factors from the term structure of interest rates and the procedure is compared with other very popular techniques: principal component analysis and Nelson-Siegel model. The NNMF approximation for the term structure of interest rates is better in terms of fitting. From a practitioner point of view, the NNMF factors and factor loadings obtained possess straightforward financial interpretations due to their non-negativeness.

7 Hellinton Takada and Julio Stern, Information Criterion for Selection of Ubiquitous Factors

Factor analysis is a statistical procedure to describe observed data in terms of unobserved variables called factors. Naturally, it is necessary to determine the number of factors to represent the system and there are several existent criteria to deal with the tradeoff between reduction of approximation error and avoidance of overparameterization. However, given the factors there is a lack of an approach to verify if they are really equally inherent to the entire data. In this paper, the term ubiquitous factors is coined to describe such equally omnipresent factors and it is proposed an information criterion to fill the existent blank. Additionally, it is also shown the possibility to use the criterion to compare ubiquity of factors from two different techniques: principal component analysis and non-negative matrix factorization. Finally, the proposed criterion is extended to identify factors more suitable to describe only a partition of the data.

15 Alexis Decurninge and Frédéric Barbaresco,Robust Burg Estimation of stationary autoregressive mixtures covariance

For a Gaussian autoregressive process, Burg methods are often used in case of stationarity for their efficiency even when few samples are available. Unfortunately, if we directly apply these methods to estimate the common covariance matrix of N vectors coming from a non-Gaussian distribution, the efficiency will strongly decrease. In this article, we adapt these methods to mixtures of Gaussian autoregressive vectors by changing the energy functional to minimize in the Burg algorithm. We will propose estimators both independent of the non-Gaussian amplitude of the considered random vectors and robust to heavy contamination.

17 Keiko Uohashi, Harmonic maps relative to α-connections on statistical manifolds

Harmonic maps are important objects in certain branches of geometry and physics. In this paper, we study harmonic maps relative to α-connections, *but not necessarily standard harmonic maps.*

The standard harmonic map is defined by the first variation of the energy functional of the map. A harmonic map relative to an α-connection is defined by an equation similar to a first variational equation, though it is not induced by the first variation of the energy functional. We first recall definitions of harmonic maps relative to α-connections explained in GSI2013. Next, we show properties of equations which defines harmonic maps relative to α-connections, and compare them with integrability of the Euler-Lagrange equations on standard harmonic maps.

18 Soumia Sid Ahmed, Zoubeida Messali, Abdeldjalil Ouhabi, Sylvain Trepout, Cedric Messaoudi and Segio Marco Non Parametric Denoising Methods Based on Wavelets : Application to Electron Microscopy Images in Low Time Exposure

The 3D reconstruction of the Cryo-Transmission Electron Microscopy (cryo-TEM) and Energy Filtering TEM images hampered by the noisy nature of these images, so that their alignment becomes so difficult. This noise refers to the collision between the frozen hydrated biological samples and the electrons beam, where the specimen is exposed to the radiation with a high exposure time. This sensitivity to the electrons beam led specialists to obtain the specimen projection images at very low exposure time, which resulting the emergence of a new problem, an extremely low signal-to-noise ratio (SNR). This paper investigates the problem of TEM images denoising when they are acquired at very low exposure time. So, our main objective is to enhance the quality of TEM images to improve the alignment process which will in turn improve the three dimensional tomography reconstructions. We have done multiple tests on special TEM images acquired at different time exposure 0.5s, 0.2s, 0.1s and 1s (i.e. with different values of SNR)) and equipped by golding beads for helping us in the assessment step. We herein, propose two structures, based on five different denoising methods, to combine the multiple noisy TEM images copies. Namely, the five different methods are Soft, Semi-Soft and the Hard as Wavelet-Thresholding methods, Bilateral Filter as a non-linear technique able to maintain the edges neatly, and the Bayesian approach in the wavelet domain, in which context modeling is used to estimate the parameter for each coefficient. To ensure getting a high signal-to -noise ratio, we have guaranteed that we are using the appropriate wavelet family at the appropriate level. So we have chosen "sym8" wavelet at level 3 as the most appropriate parameter with what is required. Whereas, for the bilateral filtering many tests are done in order to determine the proper filter parameters represented by the size of the filter, the range parameter and the spatial parameter. The experiments reported in this paper demonstrate the best performance of the Bilateral Filtering and the Bayesian approaches In terms of improving the SNR out and the image quality, so that there was no big change in the golden beads diameter compared to the thresholding methods where the soft method was the best choice. Taken together, these results suggest that the Bayesian process has a potential to outperform all previous methods, where in the multi copy structure it gave us the best SNR out without change the golden beads diameter when we use a few number of TEM images which means that the Bayesian approaches can give us an enhanced average image without needing a huge amount of copies.

22 Youssef Bennani, Luc Pronzato and Maria-Joao Rendas, Most Likely Maximum Entropy for Population Analysis: a case study in decompression sickness prevention

The paper proposes a new estimator for non-parametric density estimation from region censored observations in the context of population studies, where standard Maximum Likelihood is known to be affected by over-fitting and non-uniqueness problems, combining the Maximum Entropy and Maximum Likelihood criteria.

By combining the two criteria we propose a novel density estimator that is able to overcome the singularities of the Maximum Likelihood estimator while maintaining a good fit to the observed data, and illustrate its behavior in real data (hyperbaric diving).

The density estimation problem is motivated by a problem of population analysis: we are interested in the distribution pi_(theta) of the biophysical parameters theta of a mathematical model [1] for the instantaneous volume of micro-bubbles flowing through the right ventricle of a diver's heart when executing a decompression profile P:

The instantaneous gas volume (B) is observed only through measurements of bubble grades G, strongly quantified versions of the peak volume b(theta, P) =max_t B(t, theta, {P(u)}_{u <= t}).

The problem is thus region-censored: all parameter values in the regions R_i

= { theta \in Theta : b(theta,P_i) \in [tau_{G_i}, tau_{G_i+1} [}

yield the same observed grade G_i for profile P_i.

Several facts are known about the NPMLE for interval-censored observations: (i) its support is confined to a finite number of disjoint intervals (the so called ``elementary regions''); (ii) all distributions that put the same probability mass in these intervals have the same likelihood; (iii) there is in general no unique assignment of probabilities to the elementary regions that maximizes the likelihood.

Together, these facts imply that the NPMLE will frequently exhibit a singular behavior, in the sense that its mass is concentrated in a subset of Theta of small Lebesgue measure. This may lead to dangerous biases in the context of risk assessment, by not taking into consideration the presence of individuals for which risk can be large.

In this paper, we remove ambiguity in the estimation of pi_\theta by relying in the principle of Maximum Entropy (maxent, for short) [2], that finds the most un-informative density that can match the observed data. In our case, the constraints defining the Maximum Entropy solution are the empirical grade distributions observed in the set of executed diving profiles. Note that these constraints are not observed simultaneously, but derived from independent samples of the population.

While maxent has been frequently used for density estimation from joint observation of empirical moments of a set of features, its use for region-censored data arising from strongly quantified data from independent observations is, as far as we know, novel. In particular we show that the equivalence between regularized maxent and penalized Maximum Likelihood in the exponential family [3,4] is lost in our case. This leads us to formulate a new estimation criterion, finding the most likely constrained maxent density. We compare the proposed estimator to the NPMLE and to the best fitting maxent solutions in real data from hyperbaric diving, showing that the resulting distribution is a best candidate than NPMLE or maxent alone for the distribution of biological parameters in a given population.

[1] Julien Hugon, Vers une modelisation biophysique de la decompression, Ph.D. thesis, Aix Marseille 2, 2010.

[2] E. T. Jaynes, "Information theory and statistical mechanics," Physics Reviews, vol. 106, pp. 620630, 1957.

[3] M. Dudik, Maximum entropy density estimation and modeling geographic distributions of species, Ph.D. thesis, Princeton University, Department of Computer Science, 2007.

[4] M. Dudik, M. Phillips, and R. Schapire, "Performance guarantees for regularized maximum entropy density estimation," Proc. 7th Annual Conf. on Comp. Learning Theory, 2004.

46 Olivier Schwander, José Picheral, Nicolas Gac, Daniel Blacodon and Ali Mohammad-Djafari, Aero-accoustics source separation with sparsity inducing priors in the frequency domain

We study aero-accoustics source separation from observations taken in a wind tunnel. Our problem is particular since data are not measured by a sensor array outside the air flow, like in most experimental setups, but by sensors directly mounted on the observed object. This setup bears an important physical prior knowledge about the position of the sources but brings new difficulties: in particular, the sensors are sensitive to the aero-accoustics noise produced by the air flow.

We focus here on the study of a plane wing: the main source of interest in our setup is the signal emitted by the trailing edge of the wing.

This source is of particular interest for aero-accoustic specialists since the physical phenomenon producing this signal is not completely understood: producing a temporal representation of the signal would be of great interest for a better understanding of the wing behavior.

Fortunately, some physical information is known: the signal is supposed to be stationary, meaning we can work on small sub-block of the sample, and to have a sparse and unimodal frequency distribution.

To exploit this physical knowledge we use a classical Bayesian approach for source separation: the physical prior on the frequency distribution of the source of interest is translated into a sparsity inducing prior.

The first possible approach is to use, in the frequency domain, a heavy-tailed distribution like the Laplace law or the generalized Gaussian distribution: these priors induce sparsity of the frequencies but do not constraint the distribution to be unimodal. The second possible approach is to model the temporal signal by an auto-regressive model of order 2: the dual of this prior in the frequency domain is a unimodal distribution whose position and variance can be manually set (or learned as hyper-parameters). The optimization problem of the source separation is solved with a joint maximization approach.

Since the physical phenomenon is not fully known, we do not have any ground truth or simulation data and we thus need to work directly on real data from the wind tunnel. In order to evaluate our results, we exploit the physical knowledge on the source: due to its stationary nature, we can check that the separation results are coherent on each *sub-block of the observations and we can compare the frequency peak to frequency information got from source localization methods.*

56 Ramandeep Johal, Renuka Rai and Guenter Mahler, Bounds on Thermal Efficiency from Inference

There are many physical situations where one has to make inference from incomplete available information. Statistical inference has emerged as a powerful tool in recent years with applications in many diverse areas of research, such as particle physics, cosmology and astrophysics, machine learning and so on. In this

presentation, we apply the problem of inference to the process of work extraction from two finite, constant heat capacity reservoirs, when the final thermodynamic coordinates of the process are not fully specified. First we discuss the case when the values of final temperatures have been specified but it is not known as to which reservoir a specific value refers to. Given a pair of possible temperature values, we find the estimate by invoking Laplace's principle of insufficient reason which assigns equal probabilities for the occurance of different values, in the absence of any evidence to the contrary. The estimates for thermal efficiency indicate that the uncertainty

about the specific reservoir labels on the temperatures, reduces the maximal efficiency below the Carnot value, its minimum value being the well known Curzon-Ahlborn value. Thus we obtain the latter efficiency from a novel perspective. We map the inferred properties of the incomplete model to a model with complete information but with an additional source of thermodynamic irreversibility. In the second part, we make an average estimate of the efficiency, assuming this time that even the values of the temperatures are not available. This may be achieved by averaging over a uniform prior distribution. It is found that if the labels are known with certainty but just the values of temperature are unknown, then in the near-equilibrium limit the efficiency scales as 1/2 of Carnot value. On the other hand, if there is maximal uncertainty in the labels alongwith the uncertainty in value, then the average estimate for efficiency drops to 1/3 of Carnot value. The connection of these inferences with certain optimal results from the finite-time. Models of heat engines will be pointed out.

References: 1. R.S. Johal, R. Rai, G. Mahler, LANL arXiv 1305.6278, v1 (2013).

60 Yasmín Navarrete, Sergio Davis and Gonzalo Gutiérrez, Maximum entropy modelling of opinions in social groups

In this work we develop an agent-based model to study how the opinions of a group of individuals determine their spatial distribution and connectivity. The interaction between agents is derived from Maximum Entropy reasoning and is described by a Hamiltonian in which agents are allowed to move freely without an underlying lattice (the average network topology connecting them is determined from the parameters). Control parameters emerge as Lagrange multipliers of the maximum entropy problem, and they can be associated with the level of consequence between the personal beliefs and external opinions, and the tendency to socialize with peers of similar or opposing views. Our model presents both first and second-order phase transitions, depending on the ratio between the internal consequence and the interaction with others.

62 Mina Aminghafari and Maryam Hematti, Comparing Entropy and Energy Goodness-of-fit test for Lévy Distribution

The main restriction in a parametric method is the assumption of knowing population distribution. However, there is no parametric test for testing population distribution. Goodness-of-fit tests (GFT) are introduced as a non-parametric tool as a solution of this problem. GFT has a long history. Vasicek (1976), for the first time, proposed a GFT based on entropy and it extended by Gokhale (1983) and Ebrahimi et al. (1992) in two different directions. After these contributions many papers dedicate to the application or improvement of GFTs. Behind of the well the known GFTs, recently a GFT based on energy statistics (e.g. Székely and Rizzo (2013)) is proposed for testing heavy tailed Lévy stable distributions, Yang (2012). In this work, we introduced a modified entropy base GFT for Lévy distribution. We compare the power of the proposed test with the entropy and energy GFTs through a Monte Carlo simulation.

63 Leila Gharsalli, Hacheme Ayasso, Bernard Duchêne and Ali Mohammad-Djafari, Variational Bayesian Approach with a heavy-tailed prior distribution for solving a non-linear inverse scattering problem

We consider a nonlinear inverse scattering problem where the goal is to reconstruct an image of an unknown object from measurements of the scattered field that results from its interaction with a known wave in the microwave frequency range. The forward modeling of the wave-object interaction is tackled through a domain integral representation of the electric field in a 2D-TM configuration. The inverse problem is solved in a Bayesian framework where the prior information is introduced via a heavy-tailed distribution [1] as sparse prior information. In fact, the sought object or the image to be reconstructed is supposed to be composed of homogeneous areas. This implies that the image gradient must be sparse. A Variational Bayesian Approximation (VBA) technique [2] is then applied to compute the posterior estimators and reconstruct the object.

[1] A. Mohammad-Djafari: Bayesian approach with prior models which enforce sparsity in signal and image processing. EURASIP journal on Advances in Signal processing 2012, Special issue on Sparse Signal processing.

[2] V. Smidl, A. Quinn: The Variational Bayes Method in Signal Processing,(Signals and Communication technology). Secausus, NJ, USA: Springer-Verlag New York, Inc. 2006.

80 Abdelbassit Boualem, Meryem Jabloun, Philippe Ravier, Marie Naiim and Alain Jalocha, Assessment of two MCMC algorithms convergence for Bayesian estimation of the particle size distribution from multiangle dynamic light scattering measurements

Dynamic light scattering (DLS) is a standard technique widely used to assess the particle size distribution (PSD) of a sample composed of particles dispersed in a liquid [1]. The size of particles ranges from a few nanometers to several micrometers. The increasing success of the DLS technique comes from the fact that it is easy to handle, non-destructive, fast, and requires no calibration process. Several instruments are commercially available, among them the Nano DS from CILAS. This technique is based on the analysis of the temporal fluctuations of light intensity scattered at a given *angle by the dispersed particles illuminated by a Laser beam. Retrieving the PSD information from DLS measurements (time autocorrelation function (ACF)) that are generally corrupted by noise is known as a high ill-posed inverse problem [2].*

Methods based on Bayesian inference have been recently investigated to solve this inverse problem [3-5]. In [6], we have proposed a new Bayesian inference method applied directly to the multiangle DLS measurements to improve multimodal PSDs estimation. The posterior probability density of interest is sampled using MCMC Metropolis-Hasting algorithm one variable at time. Figure 1 shows improvement of multimodal PSD estimation using the proposed method.

Since MCMC methods are associated in practice with unknown rates of convergence, the aim of the present work is to assess the convergence of the method proposed in [6] and of some other MCMC algorithms that can be used to sample the posterior probability density of interest [7]. To this end, we propose to use the simulation method recently proposed by the authors of [8] and which method help assess the MCMC algorithms efficiency. This simulation method is based on the evaluation of the Kullback divergence criterion requiring an estimate of the entropy of the algorithm successive densities.

81 Isaac Almasi and Adel Mohammadpour, Bayesian Reconstruction in Lévy Distribution

In this paper, we consider the situation in which a middle part of the order statistics is lost or removed from the experiment, while the underlying distribution is Levy distribution. The maximum-likelihood(MLR) and

Bayesian reconstruction(BR) for Levy distribution is obtained.Also, numerical example and Monte Carlo simulation study of the L´evy distribution are given to illustrate all the reconstruction methods discussed

in this work. To evaluate the estimators, we compute Mean Square Reconstruction Errors for the MLR and BR reconstructions. Finally, we conclude BR is better than MLR.

65 Ben Placek and Kevin Knuth, A Bayesian Analysis Of Kepler-2b Using The EXONEST Algorithm

The study of exoplanets (planets orbiting other stars) is revolutionizing the way we view our universe. High-precision photometric data provided by the Kepler Space Telescope (Kepler) enables not only the detection of such planets, but also their characterization. This presents a unique opportunity to apply Bayesian methods to better characterize the multitude of previously confirmed exoplanets. This paper focuses on applying the EXONEST algorithm to characterize the transiting short-period-hot-Jupiter, Kepler-2b. EXONEST evaluates a suite of exoplanet photometric models by applying Bayesian Model Selection, which is implemented with the MultiNest algorithm. These models take into account planetary effects, such as reflected light and thermal emissions, as well as the effect of the planetary motion on the host star, such as Doppler beaming, or boosting, of light from the reflex motion of the host star, and photometric variations due to the planet-induced ellipsoidal shape of the host star. By calculating model evidences, one can determine which model best describes the observed data, thus *identifying which effects dominate the planetary system. Presented are parameter estimates and model evidences for Kepler-2b.*

77 Mohsen Salehi, Adel Mohammadpour and Mina Aminghafari ,TFBS Prediction with Stochastic Differential Equation and Time Series

In molecular biology and genetics, the transcription factor binding sites (TFBS) are the regions on DNA caused a gen is expressed.

Prediction of these regions is crucial for them. Several studies have been done, such as applying position weight matrix (PWM) and logistic regression (LR), to distinguish true binding regions from random ones. We considered the Chromosome 1 and tried to use the time series and stochastic differential equation to improve the predictions. We were interested to use the distance of binding sites from each other to predict TFBS regions. In chromosome 1, we dealt with two types binding site, 5' to 3' and 3' to 5' binding sites. We plotted them and realized that the patterns of them are different, so we considered three features for our study.

At first, we worked with total of binding sites, regardless of type of them, then we did on 5' to 3' binding sites and finally, on 3' to 5' binding sites. We used two approach, time series (TS) and stochastic differential equation (SDE), to find better predictions of binding sites rather than applying PWM and LR. In the time series method, we used of Fourier series to find a pattern on distances of BS's, then in SDE method, we considered the distances of BS's as a stochastic process to predict them. We compared our results to those using PWM and LR. The results show that SDE method forecast TFBS's better than TS and applying these two method can predict TFBS regions more successfully than PWM and LR.

83 Mahdi Teimouri, Saeid Rezakhah and Adel Mohammadpour, Entropy-based goodness-of-fit test for positive stable distribution

A goodness-of-fit test for positive stable distribution is proposed. For this mean, the Kullback–Leibler distance measure, as a basic tool in entropy theory, is considered. A simulation study is performed to compare the performance of proposed method and another one that works based on a characterization of the positive stable distribution.

88 Ali Ghaderi, On coarse graining of information and its application to pattern recognition

In pattern recognition one is concern with finding regularities in data and classifying them into different categories [1]. Objects in the same category are more similar to each other than those in other categories. However, often the notion of category cannot be precisely defined. Therefore, categories in these cases are defined as collection of objects which are likely to share the same properties. One common approach to such problems is based on the so called *finite mixture models* [2]. More precisely, suppose that X is a random variable which takes values in a sample space $\mathcal{X}$, and its distribution is represented by the probability density function $p\left(x \mid \psi\right)$.

Then

$$p\left(x \middle| \boldsymbol{\psi}\right) = \sum_{j=1}^{k} \pi_j f_j\left(x \middle| \boldsymbol{\theta}_j\right), \; x \in \mathcal{X}$$

where

$$\sum_{j=1}^{k} \pi_j = 1, \; \pi_j \geq 0$$

and

$$\int_{\mathcal{X}} f_j\left(x \middle| \boldsymbol{\theta}_j\right) dx = 1, \; f_j\left(x \middle| \boldsymbol{\theta}_j\right) \geq 0$$

In such a case, one says that X has a finite mixture distribution and that $p\left(x \middle| \boldsymbol{\psi}\right)$ is a finite mixture density function. The parameters π_j are called *mixing weights* and f_j the *component densities* of the mixture.

In the context of pattern recognition, k is the number of categories and f_j is the density function describing the distribution of members of the category j. Technically, once f_j are specified, determining $(\pi_j, \boldsymbol{\theta}_j)$ becomes a standard problem in statistical inference. We argue that in order to be able to specify f_j one has to be able to relate properties of each member of a category to the properties of the whole category itself. We show how in some cases this can be achieved through coarse graining of information within each category and how it can be used to derive the functional form of f_j. The arguments will be elucidated with examples.

Key Words: Mixture probability, Coarse graining, Maximum Entropy, Pattern Recognition.

References:

[1] C. M. Bishop, Pattern Recognition and Machine Learning, Springer Science+Business Media, 2006.

[1] D. M. Titterington, A. F. M. Smith, U. E. Makov, Statistical Analysis of Finite Mixture Distributions, John Wiley & Sons Ltd., 1985.

90 Veroni Jayawardana, Adom Giffin and Sumona Mondal, Bayesian analysis of factors associated with Fibromyalga syndrome subjects

In a previous study, factors contributing to movement-related fear were assessed for subjects with Fibromyalgia (FM). The data for the study were collected by a national internet survey of community-based individuals with FM. In this study the dependent variable was Revised Fibromyalgia Impact Questionnaire (FIQR) and independent variables were Activities-Speci..c Balance Confidence Scale (ABC), Primary Care Post-Traumatic Stress Disorder Screen (PC-PTSD), Tampa Scale of Kinesiophobia (TSK), a joint hypermobility syndrome screen (JHS), Vertigo Symptom Scale (VSS-SF), Obsessive-compulsive personality disorder (OCPD), pain, work status and physical activity. In this paper, we revisit this data by performing a Bayesian analysis where we introduce appropriate priors from previous work. A linear model was assumed in the previous study and this will be the case here as well.

91 Adom Giffin, Joseph Skufca and Palmer Lao, Using Bayes factors for multi-factor, biometric authentification

The use of biometrics for authentication purposes is a rapidly expanding field. It is well known that combining methods of authentication increases security. Therefore, multi-factor/multi-modal authentication systems are becoming the de facto standard.

Traditional authentication methods usually calculate the True Acceptance Rate and the False Acceptance Rate. These rates are point estimates that do not have any uncertainty associated with them. Further, multiple factors are typically fused together in an ad hoc manner. To be consistent, as well as to establish and make proper use of uncertainties, we use a Bayesian method that will update our estimates and uncertainties as new information presents itself.

Our algorithm compares competing classes (such as genuine vs. imposter) using Bayes Factors (BF). The importance of this is that we do not only accept or reject one model (class), but compare it to others to make a decision. We show using a Receiver Operating Characteristic (ROC) curve that using BF for determining class will always perform at least as well as the traditional combining of factors, such as a voting algorithm. As the uncertainty decreases, the BF result continues to exceed the traditional methods result. In addition, by using BF, not only can we fuse the data but we can suggest what biometric factor should be asked for next. Furthermore, we show how we use this algorithm to reduce redundancies in particular biometric factors such as voice.

93 Steven H. Waldrip, Robert K. Niven, Markus Abel and Michael Schlegel
MaxEnt Analysis of a Water Distribution Network in Canberra, ACT, Australia

*A maximum entropy (MaxEnt) method is developed to infer the state of a pipe flow network, for situations in which there is insufficient information to form a closed equation set. This approach substantially extends existing deterministic methods for the analysis of engineered flow networks (e.g.\ Newton's method or the Hardy Cross scheme). The network is represented as an undirected graph structure, in which the uncertainty is represented by a continuous relative entropy on the space of internal and external flow rates. The head losses (potential differences) on the network are treated as dependent variables, using specified pipe-flow resistance functions. The entropy is maximised subject to ``observable'' constraints on the mean values of certain flow rates and/or potential differences, and also ``physical'' constraints arising from the frictional properties of each pipe and from Kirchhoff's nodal and loop laws. A numerical method is developed in Matlab for solution of the integral equation system, based on multidimensional quadrature. Several non-linear resistance functions (e.g. power-law and Colebrook) are investigated, necessitating numerical solution of the implicit Lagrangian by a double iteration scheme.\\ *

The method is applied to a 1123-node, 1140-pipe water distribution network for the suburb of Torrens in the Australian Capital Territory, Australia, using network data supplied by

water authority ACTEW Corporation Limited. A number of different assumptions are explored, including various network geometric representations, prior probabilities and constraint settings, yielding useful predictions of network demand and performance. We also propose this methodology be used in conjunction with in-flow monitoring *systems, to obtain better inferences of user consumption without large investments in monitoring equipment and maintenance. \\ *

\noindent {\bf Acknowledgements:} Digital data in this work are sourced from and owned by ACTEW Corporation Limited (ABN 86 069 381 960) trading as ACTEW Water.

Wednesday, September 24th

Oral session 7 (9h00-10h15) - Quantum physics

Chaired by: S. Zozor / J.F. Bercher / F. Barbaresco

Invited talk: **R. Balian, The entropy-based quantum metric (98)**

The von Neumann entropy S(^D) generates in the space of quantum density matrices ^D the Riemannian metric $ds^2 = - d^2S$(^D), which is physically founded and which characterizes the amount of quantum information lost by mixing ^D and ^D + d^D. A rich geometric structure is thereby implemented in quantum mechanics. It includes a canonical mapping between the spaces of states and of observables, which involves the Legendre transform of S(^D). The Kubo scalar product is recovered within the space of observables. Applications are given to equilibrium and non-equilibrium quantum statistical mechanics. There the formalism is specialized to the relevant space of observables and to the associated reduced states issued from the maximum entropy criterion, which result from the exact states through an orthogonal projection. Von Neumann's entropy specializes into a relevant entropy. Comparison is made with other metrics. The Riemannian properties of the metric $ds^2 = - d^2S$(^D) are derived. The curvature arises from the non-Abelian nature of quantum mechanics; its general expression and its explicit form for q-bits are given.

27 Ryszard Kostecki A new maximum entropy/information geometry approach to bayesian foundations of quantum theory

I will present a new approach to bayesian foundations of quantum theory. Its novelty is both conceptual and mathematical. On the mathematical side, I will show how quantum kinematics can be reformulated in the Chencov--Amari spirit, using quantum information geometry instead of spectral theory and Hilbert spaces. Moreover, I will show that Lueders' rule of quantum state change is a special case of constrained quantum entropy maximisation, which generalises the results by Williams, Caticha, Giffin, and others, to the quantum case. These results open the direct way for construction of nonperturbative nonlinear predictive quantum theories, understood as tools for inductive inference, in a very close analogy to Jaynes' approach to foundations of statistical mechanics. Finally, I will discuss a new interpretation, epistemic intersubjectivism, which is aimed to bypass the clinch between subjectivist and ontic views on quantum states and probabilities in a different way than Fuchs' and Caticha's pragmatist-inspired bayesian approaches do.

10:30 – 12:45 Oral session 8 – Quantum physics

Chaired by: Roger Balian / J.F. Bercher / Ariel Caticha / R. Kostecki

84 Ariel Caticha, Daniel Bartolomeo and Marcel Reginatto Entropic Dynamics: from Entropy and Information Geometry 38 Marcel Reginatto The geometrical structure of quantum theory as a natural generalization of information geometry

Our subject is Entropic Dynamics, a framework which emphasizes the deep connections between the laws of physics and information. In the more standard views quantum theory has been regarded as a type of mechanics and it is natural to postulate dynamical laws described by action principles. In contrast, in the entropic view quantum theory is an application of entropic methods of inference and there is no underlying action principle. The dynamics is driven by entropy as constrained by the appropriate relevant information—it is through these constraints that the "physics" is introduced.

The entropic dynamics approach allows us to see familiar notions such as time, mass, the phase of the wave function, and Hilbert spaces from an unfamiliarly fresh perspective. To begin, since incomplete information, uncertainty, and probabilities are the norm—indeterminism demands no explanation. Instead what requires an explanation—but this is much easier to achieve—is the determinism that characterizes the classical limit. The notion of time is introduced as a book-keeping device to keep track of changes and, although quantum mechanics is time reversible, time itself is not, and an arrow of time emerges naturally.

In this paper the derivation of quantum theory as a form of entropic dynamics is strengthened by introducing concepts of information geometry and by establishing its relation to Hamiltonian dynamics. The main results include: (1) We show that a non-dissipative entropic dynamics naturally leads to a Hamiltonian dynamics, including its associated symplectic geometry, and an action principle. (2) The metric of the N-particle configuration space is not postulated; it is derived from information geometry and shown to coincide with the mass tensor. (Its inverse is the diffusion tensor.) (3) The particular form of Hamiltonian that leads to quantum theory requires a so-called "quantum potential". Using information geometry we show that the quantum potential follows naturally from the mass tensor and the Fisher information.

44 Stephan Weis The MaxEnt extension of a quantum Gibbs family, convex geometry and geodesics

This talk revisits Gibbs families of probability distributions and reports new results about quantum states. In 1972 \v Cencov has defined an extension of a log-linear model of probability distributions in order to have a maximum-likelihood estimate with probability one. His ideas have advanced in mathematical statistics in the work of Barndorff-Nielsen (1978), Csisz\'ar and Mat\'u\v s ('03,'05,'08), Geiger, Meek and Sturmfels ('06) and Rauh, Kahle and Ay ('11). A log-linear model is the set of Boltzmann distributions of a vector space of Hamiltonians. Its elements are results of the inference method, formulated by Jaynes (1957), to infer a distribution from expected values of the given Hamiltonians by maximizing the entropy. The norm closure is a suitable extension of a log-linear model of finite support as it consists of all maximum-entropy *distributions and as it extends the maximum-likelihood estimate. Closures of log-linear model of infinite support are much more subtle.*

Von Neumann has introduced Gibbs states to quantum statistical mechanics in 1927 where they play a central role every since. Gibbs states solve the inverse problem to infer a state from expected values of a set of observables. An application is the reconstruction of quantum states, see for example Bu\v zek et al.\ (1999), which was also considered in the reconstruction of quantum channels by Ziman ('08).

Closure subtleties of Gibbs families begin in the quantum setting already for finite-level systems. W.\ and Knauf ('10) have shown that the maximum-entropy inference can be discontinuous. This has a convex geometric origin because the convex set of quantum states has curved and flat boundary portions, while the probability simplex has only flat boundary portions. The convex geometric origin becomes very transparent in the notion of openness of a restricted linear map, see the preprint cited below, which is also helpful to understand the geodesics in a Gibbs family. In information geometry, a (-1)-geodesic in a Gibbs family is a straight line in the expected value chart, see Amari and Nagaoka ('00). I report that the union of all (-1)-geodesics in a Gibbs family and their limit points equals the set of maximum-entropy states. This is remarkable since the analogous construction with $(+1)$-geodesics, that is straight lines in the chart of Lagrangian multipliers, can be strictly smaller. The difference has a convex geometric origin. The two constructions are equivalent for commutative observables.

Weis, S. (in press). Continuity of the maximum-entropy inference. Commun MathPhys.

arXiv:1202.3116 [quant-ph]

45 Steeve Zozor, Gustavo Martín Bosyk, Mariela Portesi, Tristan Osán and Pedro Walter Lamberti Beyond Landau--Pollak and entropic inequalities: geometric bounds imposed on uncertainties sum

In this paper we propose generalized inequalities to quantify the uncertainty principle. We deal with two observables with finite discrete spectra described by positive operator-valued measures (POVM) and with systems in mixed states. Denoting by p(A;ρ) and p(B;ρ) the probability vectors associated with observables A and B when the system is in the state ρ, we focus on relations of the form $U_\alpha(p(A;\rho))+U_\beta(p(B;\rho)) \geq B_{\alpha,\beta}(A,B)$ where U_λ is a measure of uncertainty and B is a non-trivial state-independent bound for the uncertainty sum. We propose here:

(i) an extension of the usual Landau–Pollak inequality for uncertainty measures of the form $U_f(p(A;\rho)) = f(\max_i p_i(A;\rho))$ issued from well suited metrics; our generalization comes out as a consequence of the triangle inequality. The original Landau–Pollak inequality initially proved for nondegenerate observables and pure states, appears to be the most restrictive one in terms of the maximal probabilities;

(ii) an entropic formulation for which the uncertainty measure is based on generalized entropies of Rényi or Havrda–Charvát–Tsallis type: $U_{g,\alpha}(p(A;\rho)) = g(\sum_i [p_i(A;\rho)]^\alpha)/(1-\alpha)$. Our approach is based on Schur-concavity considerations and on previously derived Landau–Pollak type inequalities.

61 Klil Neori and Philip Goyal Anyons in the Operational Formalism

The operational formalism to quantum mechanics seeks to base the theory on a firm foundation of physically well-motivated axioms [1]. It has succeeded in deriving the Feynman rules for general quantum systems. Additional elaborations have applied the same logic to the question of identical particles, confirming the so-called Symmetrization Postulate: that the only two options available are fermions and bosons~[2-3]. However, this seems to run counter to results in two-dimensional systems, which allow for anyons, particles with statistics which interpolate between Fermi-Dirac and Bose-Einstein (see [4] for a review).

In this talk we will show that the results in two dimensions can be made compatible with the operational results. That is, we will show that anyonic behavior is a result of the topology of the space in two dimensions [5], and does not depend on the particles being identical; but that nevertheless, if the particles are identical, the resulting system is still anyonic.

References:

[1] P. Goyal, K. H. Knuth and J. Skilling, "Origin of complex quantum amplitudes and Feynman's rules", Phys. Rev. A, Vol 81, 022109 (2010)

[2] K. Neori and P. Goyal, "On the origin of the quantum rules for identical particles", AIP Conference Proceedings, Vol 1551, No 1, 220-227 (2013)

[3] P. Goyal, "Informational Approach to Identical Particles in Quantum Theory", in review (2014)

[4] A. Lerda, Anyons: Quantum Mechanics of Particles with Fractional Statistics (Springer-Verlag, 1992)

[5] M. G. G. Laidlaw and C. DeWitt-Morette, "Feynman Functional Integrals for Systems of Distinguishable Particles", Phys. Rev. D, Vol 3, 1375-1378

14:00 – 15:00 Oral session 9 - Quantum entropy

Chaired by: S. Zozor / R. Balian / Roman Belavkin

64 Roman Belavkin On Variational Definition of Quantum Entropy

The entropy associated with probability measure P can be defined in at least three different ways: 1) as the expectation the Kullback-Leibler (KL) divergence of P from elementary δ-measures (in this case, it is called the expected \emph{self-information}); 2) as a negative KL-divergence of some reference measure μ from the probability measure P; 3) as the supremum of Shannon's mutual information taken over all channels such that P is the output probability, in which case it is the value of some transportation problem. In classical (i.e. commutative) probability, all three definitions lead to the same quantity, providing only different interpretations of entropy. In non-commutative (i.e. quantum) probability, however, these definitions are not equivalent. In particular, the third definition, where the supremum is taken over all entanglements of two quantum systems with P being the output state, leads to the quantity that can be twice the von Neumann entropy. It was proposed originally by V.~Belavkin and Ohya \cite{Belavkin-Ohya02:_entan} and called the \emph{proper} quantum entropy, because it allows one to define quantum conditional entropy that is always non-negative. Here we extend these ideas to define also quantum analog of *proper cross-entropy and cross-information. We show some basic properties of these quantities and discuss their corresponding variational problems.*

References:

Belavkin, V.P., Ohya, M.: Entanglement, quantum entropy and mutual information.

Royal Society of London Proceedings Series A 458(209) (January 2002)

23 Pierre Maréchal Duality for maximum entropy diffusion MRI

We derive an entropy model for diffudion MRI and show that Fenchel duality techniques make this model tractable. This generalizes a model proposed by Daniel Alexander in 2005, in which the displacement of particles is constrained to a sphere. In order to better suit the physics of the diffusion process, we propose to relax this constraint. The Kullback-Leibler relative entropy is used to measure the discrepancy between the probability to be inferred and some reference measure. The obtained optimization problem is then studied using tools from partially finite convex progamming. The solution can be computed via the unconstrained maximization of a smooth concave function, whose number of variable is merely (twice) the number of Fourier samples.

Thursday, September 25th

9:00 – 10:30 Oral session 10
Geometric Structure of Information

Chaired by: Ariel Caticha / John Skilling

Invited talk: D. Bennequin, Topological forms of information (97)

This talk will present recent joint works with Pierre Baudot, where we propose a general definition of categories of informations, and study a natural cohomology for associated information quantities, and homotopical derived notions. Entropies of Shannon, Kullback and Von Neumann, appear as first fundamental classes for classical and quantum setting respectively. The decomposition of entropy in higher mutual information functions appears as an homotopical structure, and generates a new kind of topology. Possible applications to the study of large statistical data and dynamics of neuronal systems will be mentioned

24 Ke Sun and Stéphane Marchand-Maillet Information Geometry for Semi-parametric and Supervised Density Estimation

We investigate a modern approach to kernel density estimation where the kernel function varies from point to point. Density estimation in the input space means to find *a set of coordinates on a statistical manifold. This novel perspective helps to combine efforts from information geometry and machine learning to spawn a family of density estimators. We present example models with demonstrative experiments. We discuss the principle and theory of such density estimation.*

10:45 – 12:45 Oral session 11 Learning, Information, Divergence

Chaired by: Frank Nielson / Kevin Knuth / Fred Daum / O. Rioul

26 Jean-François Bercher, Valérie Girardin, Justine Lequesne and Philippe Regnault Goodness-of-fit tests based on (h,φ)-divergences and entropy differences

We consider fitting uncategorical data to a parametric family of distributions by means of tests based on (h,φ)-divergence estimates.

The class of (h,φ)-divergences, introduced in Salicr\'u {\it et al.}~(1993), includes the well-known classes of ϕ-divergences, of Bregman divergences and of distortion measures. The most classic are Kullback-Leibler, R\'enyi and Tsallis divergences.

Most of (h,φ)-divergences are associated to (h,φ)-entropy functionals, e.g., Kullback-Leibler divergence to Shannon entropy.

Distributions maximizing (h,φ)-entropies under moment constraints are involved in numerous applications and are also of theoretic interest. Besides the family of exponential distributions maximizing Shannon entropy, see, e.g., Bercher~(2014) for an overview of various information inequalities involving the so-called q-Gaussian distributions, i.e., distributions maximizing R\'enyi (or Tsallis) entropy under variance constraints.

For distributions maximizing Shannon or R\'enyi entropy under moment constraints, the related divergence is well known to reduce to an entropy difference. Then estimating divergence reduces to estimating entropy; see Girardin and Lequesne~(2013a, 2013b).

A commonly used non-parametric procedure for estimating entropy is the nearest neighbors method; see Vasicek~(1976) for Shannon entropy and Leonenko {\it et al.} (2008) for R\'enyi entropy. Vasicek (1976) deduced a test of normality whose statistics involves Shannon entropy difference, thus opening the way to numerous authors who adapted or extended the procedure to obtain goodness-of-fit tests for various sub-families of exponential distributions.

Recently, Girardin and Lequesne (2013b) considered goodness-of-fit tests for q-Gaussian distributions (among which the non-standard Student distribution arises as a meaningful example) based on R\'enyi's divergence and entropy differences. Further, we will show how this methodology may extend to families of distributions maximizing other (h,φ)-entropies.

References:

J.-F. Bercher. (2014). On generalized Cram\'er-Rao inequalities and an extension of the Shannon-Fisher-Gauss setting. In New Perspectives on Stochastic Modeling and Data Analysis, J. Bozeman, V. Girardin and C. H. Skiadas (Eds), ISAST, 19-35.

V. Girardin and J. Lequesne. (2013a). Relative Entropy Versus Entropy Difference in Goodness-of-Fit Tests. Application to Pareto fitting. Pr\'epublication du LMNO-UCBN, Caen, France.

V. Girardin and J. Lequesne. (2013b). Entropy based goodness-of-fit tests. Proceedings 2d Marrakesh International Conference on Probability and Statistics.

N. Leonenko, L. Pronzato and V. Savani. (2008). A class of R\'enyi information estimators for multidimensional densities. Ann. Stat. 36(5) 2153-2182.

M. Salicru, M. L. Menéndez, D. Morales and L. Pardo. (1993). Asymptotic distribution of (h, φ)-entropies. Comm. Stat. (Theory and Methods) 22(7) 2015-2031.

O. Vasicek. (1976). A test of normality based on sample entropy. J. Roy. Statist. Soc. Ser. B 38(1) 54-59.

35 Takashi Takenouchi, Osamu Komori and Shinto Eguchi A novel boosting algorithm for multi-task learning based on the Itakuda-Saito divergence

In this paper, we propose a novel multi-task learning algorithm based on an ensemble learning method. In a framework of the multi-task learning, we assume that there are multiple related tasks (datasets) sharing common structures, and we can utilize the shared structures to improve generalization performance of predictors for multiple tasks. This framework has been successfully employed to various kind of applications such as medical diagnosis. In this paper, we consider a specific setting of the multi-task learning for binary classification problems, in which features are shared among all tasks and all tasks are targets of performance improvement. We focus on a situation that the shared structures are represented by relationship or divergence between underlying distributions associated with multiple tasks.

We tackle with this problem using a boosting method which makes it possible to adaptively learn complicated problems with low computational cost. The boosting methods are notable implementations of the ensemble learning and tries to construct a better classifier by combining weak classifiers without heavy computational cost. AdaBoost is the most popular boosting method and a lot of variation have been developed.

In this paper, we firstly reveal that AdaBoost can be derived by a sequential minimization of the Itakura Saito divergence between an empirical distribution and a pseudo measure model constructed by a classifier. The Itrakura Saito divergence is a special case of the Bregman divergence between positive measures and is frequently used in the region of signal processing. Secondly, we propose a novel boosting algorithm for the multi-task learning based on the boosting method. We utilize the Itakura Saito divergence as a discrepancy measure between pseudo models associated with tasks, and incorporate the Itakura Saito divergence as a regularizer into the AdaBoost. The proposed method can capture the shared structure, i.e., relationship between underlying distribution by considering divergence between pseudo models estimated by constructed classifiers.

We investigate statistical properties of the proposed method and discuss validity of the regularization by the Itakura Saito divergence with small experiments.

12 Alexander Zuevsky Towards automorphic to differential correspondence for vertex algebras

In these notes we propose a version of geometric correspondence between parameter spaces for projectively flat connections in vector bundles and automorphic representations of modular groups over Riemann surfaces. Principal role of vertex algebras is discussed. We then formulate a conjecture concerning an extended correspondence between categories of twisted $\mathcal D$-modules and Hecke eigensheaves both defined on the moduli stack of modular group bundles, and obtained as sheaves of conformal blocks for vertex operator algebras with formal parameters on complex algebraic curves.

82 Michel Broniatowski, Jana Jureckova and Amor Keziou Minimum risk equivariant minimum divergence estimates for moment condition models

The aim of the talk is to investigate the ..nite-sample optimality property estimation in the context of semiparametric model built from moment conditions. We will discuss the problem of constructing minimum risk equivariant estimates (MRE, or Pitman estimators) for the parameter of the model, as well as the problem of the numerical calculation of these estimates.

14:00 – 15:30 Oral session 12 – Bayesian learning

Chaired by: Stephan Weis */ F. Nielsen / Ariel Caticha*

78 Frank Nielsen On learning statistical mixtures maximizing the complete likelihood

Learning efficiently finite statistical mixtures is one of the most fundamental problems in signal processing with paramount applications.

In the statistics community, the prevalent method is to start from a ``good'' initialization of the mixture parameters and to optimize locally the incomplete likelihood using

the expectation-maximization algorithm~\cite{bregmankmeans-2005}.

This yields a soft clustering algorithm converging asymptotically to a local maximum.

Recently, the theoretical computer science community investigated the polynomial learnability of mixtures~\cite{PolynomialGMM-2010,IsoGMM-2013} by focusing on attaining the global optimum provided some weak conditions on component separability.

In this talk, we consider learning mixtures by maximizing the {\em complete} likelihood~\cite{kmle,kmle-Wishart}, and show that it is equivalent to solving a sequence of hard k-means-type clustering problems.

Furthermore, we show how to use dynamic programming to learn univariate mixtures~\cite{intervalclustering-2014} and report a sufficient condition for optimality *(including uni-order exponential families like Rayleigh or Poisson families, or location families like Cauchy or Laplacian families).*

We discuss on several extensions, present experiments and mention several open problems.

93 U. von Toussaint, Robust phase estimation for signals with a low signal-to-noise-ratio

*A maximum entropy (MaxEnt) method is developed to infer the state of a pipe flow network, for situations in which there is insufficient information to form a closed equation set. This approach substantially extends existing deterministic methods for the analysis of engineered flow networks (e.g.\ Newton's method or the Hardy Cross scheme). The network is represented as an undirected graph structure, in which the uncertainty is represented by a continuous relative entropy on the space of internal and external flow rates. The head losses (potential differences) on the network are treated as dependent variables, using specified pipe-flow resistance functions. The entropy is maximised subject to ``observable" constraints on the mean values of certain flow rates and/or potential differences, and also ``physical" constraints arising from the frictional properties of each pipe and from Kirchhoff's nodal and loop laws. A numerical method is developed in Matlab for solution of the integral equation system, based on multidimensional quadrature. Several non-linear resistance functions (e.g. power-law and Colebrook) are investigated, necessitating numerical solution of the implicit Lagrangian by a double iteration scheme.\\ *

*The method is applied to a 1123-node, 1140-pipe water distribution network for the suburb of Torrens in the Australian Capital Territory, Australia, using network data supplied by water authority ACTEW Corporation Limited. A number of different assumptions are explored, including various network geometric representations, prior probabilities and constraint settings, yielding useful predictions of network demand and performance. We also propose this methodology be used in conjunction with in-flow monitoring systems, to obtain better inferences of user consumption without large investments in monitoring equipment and maintenance. \\ *

\noindent {\bf Acknowledgements:} Digital data in this work are sourced from and owned by ACTEW Corporation Limited (ABN 86 069 381 960) trading as ACTEW Water.

73 Philippe Cuvillier Time-coherency of Bayesian priors of transient semi-Markov chains for sequential alignment

This paper proposes a novel insight to the problem of duration modeling for Information Retrieval problems where a discrete sequence of events is estimated from a time-signal using Bayesian models. Since the duration of each event is unknown, a major issue is setting the right Bayesian prior on each of them. Hidden Semi-Markov models (HSMM) allow choosing explicitly any probability distribution for the durations but learning these statistically is a non-parametric problem. In absence of huge training data sets, most algorithms rely on regularization techniques such as choosing parametric classes of distributions but the justifications of such techniques are often heuristics.

Among the numerous application domains of HMM-like paradigms, music-to-audio alignment brings two interesting properties. Firstly, a music score informs of the ordering among events. Secondly, it assigns to each event a nominal duration. For alignment tasks the Markov models conveniently model the ordering with transient chains. But the modeling of these nominal durations is a crucial and undermined problematic. This work investigates the relationship of this prior information of duration with the Bayesian priors of a HSMM. Theoretical insights are obtained through the study of the prior state probability of transient semi-Markov chains. Whereas ergodic chain and their convergence to an equilibrium probability are well studied, transient chains constitute an undermined case but of prime importance for real-time inference on HSMM.

On the first hand we prove that the non-asymptotical evolution of the state probability has some particular behaviors if the Bayesian priors fulfill several precise conditions, derived from statistical properties like the hazard rate and the tail decay. Then we say that a model is time-coherent if the evolution of the state probability respects the information of ordering and nominal lengths. This leads to several prescriptions on the design of HSMM Bayesian priors. On the other hand we get further prescriptions by comparing the Bayesian priors associated to different nominal lengths. This real-valued parameter comes with a natural ordering; we explain why this ordering among parameters is coherently modeled by some specific stochastic orderings among distributions that are standard in statistics. We conclude by demonstrating the practical consequences of these properties with an experiment of real-time audio-to-score alignment with the HSMM-based software Antescofo developed at Ircam.

15:45 – 18:30 Poster session 3

Chaired by: Mathieu Kowalski / *Aurélia Fraisse* / Olivier Rioul / Sueli Costa

5 Justine Lequesne and Valérie Girardin, Analysis of information into marginal effects

In this poster, we propose to construct a goodness-of-fit test for the Student distribution based on the alpha-relative entropy. The Student distribution maximizes the Rényi entropy among all distributions having a fixed variance. The proposed test is premised on a Pythagorean equality, involving the maximum entropy property of the Student distribution, which reduces to construct a test with the Rényi entropy difference.

The Rényi entropy and alpha-relative entropy are both generalizations of the Shannon entropy and the Kullback-Leibler divergence.

The proposed test thus generalizes the goodness-of-fit test of normality based on the Shannon entropy introduced by Vasicek (1976), and then extended for the exponential family ; see Girardin and Lequesne (2012) for power analysis and asymptotic properties.

In the same way, we can construct tests based on Rényi entropy for the alpha-exponential family, which contains many distributions such as q-gaussian distributions well used in statistical mechanics.

The construction of tests based on Rényi entropy requires non-parametric estimators of entropy. Vasicek proposed an estimator of the Shannon entropy based on sample *spacings with good asymptotic properties. This estimator has been also constructed for the R\'enyi entropy by Wachowiak (2005). A commonly used method for estimating Shannon or Rényi entropy is the nearest neighbors method ; see Leonenko (2008) for construction and properties. These two methods of estimation will be developped in the poster.*

13 Marco Congedo and Alexandre Barachant, A Special Form of SPD Matrix for Interpretation and Visualization of Data Manipulated with Riemannian Geometry

Currently the Riemannian geometry of symmetric positive definite (SPD) matrices is

gaining momentum as a powerful tool in a wide range of engineering applications such as image, radar and biomedical data signal processing. If the data is not natively represented in the form of SPD matrices, typically we may summarize them in such form by estimating covariance matrices of the data. However once we manipulate such covariance matrices on the Riemannian manifold we lose the representation in the original data space. For instance, we can evaluate the geometric mean of a set of covariance matrices, but not the geometric mean of the data generating the covariance matrices, the space of interest in which the geometric mean can be interpreted. As a consequence, Riemannian information geometry is often perceived by non-experts as a "black-box" tool and this perception prevents a wider adoption in the scientific community. Hereby we show that this limitation can be overcome by constructing a special form of SPD matrix embedding both the covariance structure of the data and the data itself. Incidentally, whenever the original data can be represented in the form of a generic data matrix (not even square), this special SPD matrix enables an exhaustive and unique description of the data up to second-order statistics. This is achieved embedding the covariance structure of both the rows and columns of the data matrix, allowing naturally a wide range of possible applications and bringing us over and above just an interpretability issue.We demonstrate the method by manipulating satellite images (pansharpening) and event-related potentials (ERPs) of an electroencephalography brain-computer interface (BCI) study. The first example illustrates the effect of moving along geodesics in the original data space

and the second provides a novel estimation of ERP average (geometric mean), showing that, in contrast to the usual arithmetic mean, this estimation is robust to outliers. In conclusion, we are able to show that the Riemannian concepts of distance, geometric mean, moving along a geodesic, etc. can be readily transposed into a generic data space, whatever this data space represents.

19 Jean-François Degurse, Information geometry for radar detection in heterogeneous environments

Space-Time Adaptive Processing (STAP) performs two-dimensional space and time adaptive filtering where different space channels are combined at different times. In the context of radar signal processing, the aim of STAP is to remove ground clutter returns, in order to enhance slow moving target detection. Filter's weights are adaptively estimated from training data in the neighborhood of the range cell of interest, called cell under test (CUT). The estimation of these weights is always deduced, more or less directly, from an estimation of the covariance matrices of the received signal, which is the key quantity in the process of adaptation. This STAP method is usually referred to as the sample matrix inversion (SMI). One main consideration goes into the choice of the training covariance matrices: how many and which matrices share the *same statistics with the data sample to which the weights are to be applied. On one hand, the statistics of the clutter often change very quickly and, on the other hand, we want to use as many matrices as possible to obtain a good estimate of the covariance matrix that minimizes the estimation loss. The traditional way to answer these questions is to use the classical Euclidean distance which relies on a power selection criterion. One can clearly see that this method, which works only on the signal power of the covariance matrix, doesn't take advantage of the structure of the covariance matrix. We propose in these paper a new criterion based on a physical point of view and look for the minimum distance that fits to the minimization problem of the STAP filter. We show that this distance outperforms the classical Euclidian distance. We extend this distance to the Riemannian distance. When working on Hermitian positive-definite matrices, it is natural and desirable to work with the information geometry metric. We show that the distance associated to the information geometry metric performs very well in the detection of clutter non-homogeneity and we compare two processing using both the classical Euclidian and information geometry metric based the Riemannian distance. These results lead to the hypothesis that information geometry may also be used for the computation of the mean of the selected covariance matrices, enabling better performance for target detection in heterogeneous environments.*

21 Alice Le Brigant, Marc Arnaudon and Frédéric Barbaresco, Probability on spaces of curves and in the associated metric spaces via information geometry: Application to the statistical study of non-stationary Radar signatures

Today's surface radars face a new challenge: detecting swift, low-altitude targets, which are melted in strongly inhomogeneous environments (ground clutter, sea clutter). The abrupt variations of such environments, or fluctuations of the target itself in some cases, lead to the non-stationarity of the radar signal. The aim of this doctoral thesis is to define a CFAR (Constant False Alarm Rate) detector optimized for the hypothesis of a non-stationary signal.

The detector yet to define is based on a statistical analysis of the radar signal, which we study as a time-varying path in a differential manifold. We assume that the signal is locally stationary, and we represent each stationary portion by an autoregressive process, the parameters of which (or equivalently, the covariance matrix of which) we are looking to estimate. Depending on the chosen representation (coefficients of the autoregressive model or covariant matrix), the time-varying radar signal can be seen as a path in the corresponding manifold (the Poincaré disk or the space of Toeplitz matrix). We are then confronted with the study of paths (or curves) in differential manifolds, and a key point is to be able to compute the distance between two curves. To do so, we need to define a "good" Riemannian metric on those curve spaces. Such a metric enables us to compute the length of a path connecting two curves and we can define the distance between two curves as the length of the shortest path that connects them, that is a geodesic path.

We want our distance to be invariant under reparameterization, i.e. we would like the distance between two curves to be the same whatever the chosen parameterization. In other words, we want to induce a distance on the space of curves modulo reparameterization, which we call the shape space. Consequently, in our case, a "good metric" is a metric that is invariant under the action of reparameterization, so that it induces a Riemannian metric on the shape space. The metric that I am building is an *adaptation of that of Bauer et al. [1] to the space of curves in the hyperbolic space. The idea is to consider a Sobolev-type metric of order one as the pullback metric of the L2-inner product by a certain transformation. This gives a metric with the required reparameterization invariance, which induces a Riemannian metric on the shape space.*

References :

[1] M. Bauer, M. Bruveris, S. Marsland, P. W. Michor, Constructing reparameterization invariant metrics on spaces of plane curves, (2012)

[2] P. W. Michor, D. Mumford, Riemannian geometries on spaces of plane curves (2006)

[3] R. Dahlhaus, Locally stationary processes (2012)

31 Muhammed Sutcu and Ali Abbas, First Order Dependence Trees with Cumulative Residual Entropy

This paper presents a method to approximate discrete joint probability distributions using first order dependence trees and the concept of cumulative residual entropy measure. A first order dependence tree is one that each variable is conditioned on at most one variable, and there cannot be a cycle between the variables. The cumulative residual entropy measure is the entropy functional applied to the survival function as arguments instead of the probability. The results parallel Chow-Liu's approximation of joint probability distributions using the traditional Kullback-Leibler divergence and mutual information. We formulate the cumulative residual Kullback-Leibler (KL)-divergence and the cumulative residual mutual information in terms of cumulative the survival function. We then show that the optimal first-order dependence tree approximation of the joint distribution using the cumulative Kullback-Leibler divergence is the one with the largest sum of cumulative residual mutual information pairs. We compare the results with those of Chow-Liu using the traditional entropy, KL-divergence, and mutual information. We show that the two approximations perform almost equally but they are not same. We then explore under what conditions or what kind of decision problems that CRE based first order dependence trees preferred to use, and answer "which mutual information pairs that we choose to add to the dependence tree to approximate the most accurate first order dependence tree approximation distribution, if several mutual information between pairs of variables are equal?". We finally analyze and quantify the trade-off between additional assessments or additional variables and the accuracy of the approximation.

47 Steeve Zozor and Jean-Marc Brossier, De Bruijn identity: from Shannon entropy and Fisher information to generalized $f-$divergences and $f-$Fisher divergences

In this paper we propose a generalization of the usual deBruijn identity that links the Shannon differential entropy (or the Kullback–Leibler divergence) and the Fisher information (or the Fisher divergence) of the output of a Gaussian channel. The generalization makes use of Salicrú entropies on the one hand, and of divergences of the Csizàr class on the other hand, as generalizations of the Shannon entropy and of the Kullback–Leibler divergence respectively. The generalized deBruijn identities *induce the definition of generalized Fisher informations and generalized Fisher divergences; some of such generalizations exist in the literature. Moreover, we provide results that go beyond the Gaussian channel: we are then able to characterize a noisy channel using general measures of mutual information, both for Gaussian and nonGaussian channels.*

48 Anass Bellachehab and Jeremie Jakubowicz , Distributed consensus for metamorphic systems using a gossip algorithm for CAT(0) metric spaces

We present a novel application of distributed consensus algorithms to metamorphic systems. A metamorphic system is defined as a set of identical units that can self-assemble to form a rigid structure. For instance, one can think of a molecule composed of atoms. The system can change its shape in order to adapt to different environments via reconfiguration of its constituting units. There are two constraints : (i) units cannot overlap during reconfiguration (ii) overall connectivity should always be maintained. Following Abrams, Ghrist and Peterson [1,4], a system is represented as a set of occupied vertices on a given graph G = (V,E). The vertices V of this graph represent the possible locations of the units, while the edges E of the graph represent the possibility of a movement between two locations. The nonoverlap constraint implies that one of the edge endpoints should not be occupied.

We assume in this work that several metamorphic systems form a network; two systems are connected whenever they are able to communicate with each other. The aim is to synchronize all the systems that compose the network, i.e. all the systems should have the same configuration. Again, building on ideas from Abrams, Ghrist and Peterson [1, 4] and Ardila, Baker and Yatchak [2], we represent a given system configuration as a point in a cubical complex, a k-cube of which represents k commutative movements – i.e. movements that are non-overlapping whatever their order. This cubical complex happens to be locally CAT(0) and globally CAT(0) in some cases of interest. Following previous work on gossip algorithms in metric spaces [3], we are able to :

— Propose a pairwise gossip-like algorithm to achieve our goal

— Prove its convergence when the underlying cubical complex is globally CAT(0)

— Establish linear rate convergence speed

Numerical experiments are also provided and show a good match with our theoretical analysis.

References

[1] Aaron Abrams and Robert Ghrist. State complexes for metamorphic robots. The International Journal of Robotics Research, 23(7-8) :811–826, 2004.

[2] Federico Ardila, Tia Baker, and Rika Yatchak. Moving robots efficiently using the combinatorics of CAT(0) cubical complexes. arXiv preprint arXiv :1211.1442, 2012.

[3] Anass Bellachehab and J´er´emie Jakubowicz. Random pairwise gossip on CAT(κ) metric spaces, 2014.

[4] Robert Ghrist and Valerie Peterson. The geometry and topology of reconfiguration. Advances in applied mathematics, 38(3):302–323, 2007.

51 James Walsh and Kevin Knuth, Information-Based Physics, Influence and Forces

Knuth's model of direct particle-particle influence in the case of a free particle has been shown to result in the Feynman checkerboard model of the electron, from which one can derive the Dirac equation among other Fermion properties. Most interesting is the fact that the particle zig-zags back-and-forth at the speed of light in a motion Schrodinger called Zitterbewegung. In 1+1 dimensions this motion can be described using the probability that the particle is moving to the left or right at any given point in time. Here we begin to examine the behavior of the particle when it both influences others and is influenced by others. Our results suggest that the generic concept of influence is related to the classic concept of forces.

53 Mircea Dumitru and Ali Mohammad-Djafari, Estimating the periodic components of a biomedical signal through Inverse Problem modeling and Bayesian Inference with sparsity enforcing prior

The recent developments in chronobiology needs a periodic components variation analysis for the signals expressing the biological rhythms. A precise estimation of the periodic components vector is required, corresponding to short parts of such biomedical signals. The classical approaches, based on FFT methods, are inecient considering the particularities of the data (short length, sampling interval). In this paper we propose a new method for estimating the periodic components of biomedical signals, using the prior informations (reduced number of clocks in the spectrum). The proposed method considers the periodic component vector estimation as an Inverse Problem (IP) and use the Bayesian Parameter Estimation to infer over the unknowns involved in the model, i.e. the periodic components vector and the other hyperparamters. The prior information over the periodic components vector is translated as sparsity, so in the proposed model a prior law that enforces sparsity

is considered. The considered law is the Student-t distribution, viewed as a marginal distribution corresponding to a Gaussian Scale Mixture (GSM), dened via a hidden variable, representing the inverse variances, modelled as a Gamma Distribution. The hyperparameters are modelled using the conjugate priors, i.e. using Inverse Gamma Distributions. The joint posterior law of the unknown periodic components vector, hidden variable and hyperparameters is estimated via Joint Maximum A Posteriori (JMAP) and Posterior Mean (PM). For the PM estimator, the expression of the posterior law is approximated by a separable one, via the Bayesian Variational Approximation (BVA), using the Kullback-Leibler (KL) divergence. The proposed algorithm is a iterative algorithm. In the end of the article, we present the simulations results, comparing the proposed method with other existing methods. Also a comparison inside the IP approach, i.e. different priors for enforcing sparsity, is presented.

54 Li Wang, Ali Mohammad-Djafari and Nicolas Gac, Bayesian 3D X-ray Computed Tomography image reconstruction with a Scaled Gaussian Mixture prior model

In order to improve quality of 3D X-ray tomography reconstruction for Non Destructive Testing (NDT), we investigate in this paper hierarchical Bayesian methods. In NDT, useful prior information on the volume like the limited number of materials or the *presence of homogeneous area can be included in the iterative reconstruction algorithms. In hierarchical Bayesian methods, not only the volume is estimated thanks to the prior model of the volume but also the hyper parameters of this prior. This additional complexity in the reconstruction methods when applied*

to great volumes (form 5123 to 81923 voxels) involves an increasing computational cost. To reduce it, the hierarchical Bayesian methods investigated in this paper result to an algorithm acceleration thanks to the Variational Bayesian Approximation (VBA)[1] and hardware acceleration thanks to projection and back-projection operators parrallelized on many core processors like GPU[2]. In this paper, we will consider a Student-t prior on the gradient of the image implemented in a hierarchical way[3, 4, 1]. Operators H (forward or projection) and Ht (adjoint or back-projection) implanted in multi-GPU[2] have been used in this study. The different methods will be evalued on synthetic volumes "Shepp and Logan" in terms of quality and time of reconstruction. We have used several simple regularizations of order 1 and order 2. But we can also design a quadratic regularization with the method proposed by Jang Hwan Cho in [5]. Other prior models also exists[6]. Sometimes for a discrete image, we can do the segmentation and reconstruction at the same time, for example with the method DART[7, 8], then the reconstruction can be done with less projections.

57 Geert Verdoolaege, Geodesic least squares regression for scaling studies in magnetic confinement fusion

Geodesic least-squares (GLS) regression is a new parametric regression technique intended for estimating (nonlinear) relations between variables in the presence of considerable uncertainty in the measurements (regressor and response variables) and/or the regression model. The technique has been developed with a view to scaling studies in magnetic confinement fusion devices, where substantial measurement and model uncertainty may occur. The method enables arbitrary measurement distributions and regression functions, while allowing for all parameters of the chosen distribution to be modeled. Unlike the classic regression model, the conditional distribution of the dependent variable suggested by the data, which we refer to as the `observed distribution', need not be the same as the distribution proposed by the regression model, which we call the `modeled distribution'. In the terminology of the generalized linear model (GLM), our observed distribution is based on the `saturated model', which fits the data perfectly. We estimate the model parameters by minimizing the discrepancy between the observed and modeled distribution, based on the Rao geodesic distance. Hence, the method essentially solves a regression problem on the corresponding information manifold equipped with the Fisher metric. For computational efficiency a divergence function may also be deployed as the similarity measure, reminiscent of the deviance concept in the GLM.

In this contribution, we discuss the similarities and differences between GLS regression and classic methods, such as ordinary least squares and the GLM. We apply GLS to a range of synthetic data sets, demonstrating the considerably enhanced robustness compared to the classic methods in the case of outliers and questionable model assumptions. We then apply the technique by revisiting the scaling laws for the energy confinement time and the power threshold for low to high confinement in tokamaks. We again compare with established methods and finally discuss the performance of various probabilistic similarity measures.

66 Ning Chu, Ali Mohammad-Djafari, Nicolas Gac and José Picheral, A Hierarchical Variational Bayesian Approximation Approach in Robust Acoustic Imaging

In this paper, we propose to apply Variational Bayesian Approximation (VBA) in obtaining the robust acoustic imaging which in general involves an inverse problem for acoustic source localization and power reconstruction from limited noisy measurements at microphone sensors. One of the biggest challenges is that the measurement errors are often spatially non-stationary distributed at different sensors. Our contributions are that proposed VBA hierarchical framework can use the Student's-t prior with various latent parameters to model non-stationary errors, so that robust acoustic imaging would be achieved; In order to get higher spatial resolution, another Student's-t prior model can well enforce the sparse distribution of acoustic sources; Moreover, proposed VBA approach can jointly estimate unknown variables such as source powers, as well as large dimension of latent parameters, in which some of them can indicate the estimation uncertainty or confidence interval for each of estimated source powers. We also compare the proposed VBA approach and classical Joint Maximum A Posterior (JMAP) method by simulations and real data from wind tunnel experiment for the acoustic source distribution on the vehicle surface.

68 Yannis Kalaidzidis, Inna Kalaidzidis and Marino Zerial, Quantitative microcopy: markers colocalization on intracellular vesicular structures

Fusion fluorescent protein construct and fluorescently labeled antibody provide rich information about temporal-spatial distribution of proteins in the cell. Colocalization of different proteins is sign of potential interaction and/or co-regulation intracellular processes. In many biological papers the colocalization is quantified by spatial correlation of markers intensities. For cytosolic markers the correlation is reasonable way of proteins co-distribution characterization. However, for systems biology modeling the proportion of marker A that is colocalized with marker B is a key parameter. The correlation does not give this value. For the markers that reside on vesicular structures, the colocalization that is quantified as ratio of integral fluorescence intensity in color A from vesicular double-labeled structures to the total vesicular integral intensity in color A, is the value of interest. Unfortunately, the diffraction limit of light microscopy and presence of cytosolic background result in significant apparent or "random" colocalization in the highly crowded intracellular environment. In order to get colocalization value, the random colocalization has to be estimated and measured colocalization has to be accordingly corrected.

We developed two procedures of estimated random colocalization. First is based on random displacement of object in the single color plane, second is based on probabilistic model of apparent colocalization. Both procedures gave very similar results.

The correction of random colocalization in case when estimation of random colocalization gives value close to the measured colocalization was solved by Bayesian probabilistic estimation of corrected colocalization.

The correction model becomes even more complicated, when we consider triple colocalization. In this case both nominator (the triple co-localized intensities) and denominator (double colocalized intensities) are subjects of random colocalization.

The probabilistic model for correction in case of triple colocalization was developed as well.

The method of colocaliztion correction was applied to study cargo traffic through early endocytic compartments. Two markers were used for labeling partially overlapped population of early endosomes (APPL1-positive and EEA1-positive) and third marker was used for labeling cargo (EGF and LDL).

70 Kiamars Vafayi, Maximum-entropy method of moments for density estimation with large number of moments

Estimating continuous probability densities functions from measured discrete sample data has important and wide applications. It is a free-form and non-parametric estimation, and an ill-posed inverse problem, in the sense of lacking a unique solution. One method for this estimation is the maximum-entropy method of moments (MEMM) [2].

We study MEMM with order m on a finite interval of [0,1], which gives a distribution having maximum value of Gibbs-Shannon entropy, while subject to constraints that the moments up to order m of the distribution are equal to those that are calculated empirically from the sample data. The logic is that maximizing the entropy yields the least informative estimated probability density function while still agreeing with the data.

We address particularly the methods for finding solutions to the Lagrange multipliers in MEMM from the point of view of numerical stability, to understand the troublesome regions of very large number of moments (m) or too few sample data points, where the numerical schemes are sometimes unstable or do not converge to a solution. We discuss the possible sources of instability, report on numerical experiments for possible ways to mitigate them, and compare four algorithms of 1) a linear equations method, 2) newton method, 3) a hybrid algorithm, and 4) the iterative scaling method.

We also mention the new ideas presented in [2] for a Bayesian MEMM, and indicate the role played by prior information of Lagrange multipliers and that of the number of moments. We then describe the alternative and very related way of looking at the problem of determining the number of moments as a Bayesian model selection procedure.

[1] Maximum entropy in the problem of moments, Lawrence R. Mead and Nikos Papanicolaou, J. Math. Phys. 25, 2404–2417 (1984)

[2] The maximum entropy method of moments and Bayesian probability theory, Bretthorst, G. Larry, AIP Conference Proceedings, Volume 1553, pp. 3-15 (2013)

72 Kevin Knuth, The Problem of Motion: The Statistical Mechanics of Zitterbewegung

In 1930 Erwin Schrodinger showed that the eigenvalues of the velocity of a particle described by wavepacket solutions to the Dirac equation are simply +c or -c, the speed of light. This led to his coining of the term Zitterbewegung, which is German for ``trembling motion'', where all particles of matter (Fermions) zig-zag back-and-forth at only the speed of light. The result is that any finite speed less than c, including the state of rest, only makes sense as a long-term average that can be thought of as a drift velocity. In this paper, we seriously consider this idea that the observed velocities of *particles are time-averages of motion at the speed of light and demonstrate how the relativistic velocity addition rule in one spatial dimension is readily derived by considering the probabilities that a particle is observed to move either to the left or to the right.*

74 Sergio Davis, Joaquín Peralta, Yasmín Navarrete, Diego González and Gonzalo Gutiérrez, A Bayesian interpretation of first-order phase transitions

In this work we show that the formalism used in describing the thermodynamics of abrupt (or first-order) phase transitions is simply an application of maximum entropy inference. More precisely, we show that the concepts of transition temperature, latent heat and entropy difference between phases will inevitably have an equivalent in any problem of inferring the result of a yes/no question, where the initial state of knowledge is updated by new information in the form of an expectation value.

94 R. Preuss and U. von Toussaint, Bayesian uncertainty quantification for an electrostatic plasma model

Divergence functions, which are traditionally viewed as a bi-variate function on some manifold M, are here viewed as functions on the cross-manifold MxM which generate a statistical structure (a Riemannian metric plus a pair of torsion-free conjugate connections) along its diagonal manifold. Imposing compatibility conditions allow us to define a divergence function that is "proper". Further conditions can be imposed so that the cross-manifold may admit a complex representation, linking the divergence function to the "potential" on MxM. For the family of divergence functions induced by a convex function (Zhang, 2004), it is shown that they are proper and their Kahler potentials are given exactly by the inducing convex function. These results highlight the "reference-representational biduality" in Information Geometry.

Friday, September 25th

09:00 – 11:00 Oral session 13 Entropy and Information Geometry

Chaired by: John Skilling/ F. Nielsen / F. Barbaresco

39 Subrahamanian K S Moosath and Harsha K.V., F-Geometry and the Geometry Induced by a Two Point Function on a Statistical Manifold

There are many ways to introduce geometric structures on a statistical manifold S. Amari introduced a geometric structure on a statistical manifold S called alpha-geometric structure using alpha-embeddings of S into the space of random variables. Eguchi introduced a more convenient way of constructing geometric structures on a statistical manifold using a two point function called divergence functions. Using a general embedding F of S into the space of random variables Harsha and Moosath given the F-geometry on a statistical manifold. F-geometry is useful in the further development of information geometry and its applications.

*In this paper, a two point function DF is defined on a statistical manifold S using an embedding F of S into the space RX of random variables. We show that the geometry induced by DF is same as the F-geometry defined using the embedding F. Then we obtain that the geometry induced by the dual D*F of DF is same as the geometry* *induced by the dual embedding of the embedding F. Further we show that when F is the alpha-embedding, DF is same as the alpha-divergence and induced geometry is the alpha-geometry.*

49 Luigi Malagò and Giovanni Pistone, Gradient flow of the stochastic relaxation on a generic exponential family

Given an objective function f and an exponential family Ep, we assume the objective function to be bounded below, lower semicontinuous, and that the levels set near the minimum are compact. The relaxed function F is defined on the exponential family F : Ep by the expected value of f, F = E(f). If the Gibbs model based on f is part of the exponential family, it is easy to show that the minimum of the relaxed problem is equal to to the (global) minimum of the objective function. Moreover, in some cases it is possible to describe the weak limit(s) of a minimizing sequence in the exponential family as they depend on the behaviour of the objective function near the minima. In most cases the Gibbs model of the objective function will not be part of the exponential family or, this information is not available in principle because the function is given by a black box. In such a case the gradient flow of the relaxed function is considered. The gradient flow can be studied in both the exponential and the expectation parameters, the geometric gradient being represented by the natural gradient or by the ordinary gradient in the expectation parameters, respectively. We discuss some cases where it is possible to prove that the global minimum is reached by the gradient flow.

11 Harsha K V and Subrahamanian Moosath K S, Geometry of F-likelihood Estimators and F-Max-Ent Theorem

The notion of exponential family is generalized by deforming the exponential function appearing in it and this leads to the q-exponential family of probability distributions having their entropic base in Tsallis entropy. Naudts [1] studied them extensively and generalized to a large class of families of probability distributions. An information geometric foundation for the deformed exponential family is given by Amari et al.[2]. q-Pythagoras theorem and q-Max-ent theorem were discussed and a geometric proof of q-Max-ent theorem is given by Amari and Ohara [3].

In this paper we consider a family of probability distributions called F-exponential family based on the idea of (F,G)-geometry discussed in [4]. F-exponential family has got a dually flat structure which is obtained by the conformal flattening of the (F,G)-geometry. Using the canonical divergence on the F-exponential family, the F-Pythagoras theorem and the F-Projection theorem are given. Generalized notion of independence called the F-independence is introduced using the function F and its inverse. Then the F-likelihood function and the F-likelihood estimator are defined.

Further the geometry of F-likelihood estimator is discussed. Finally the F-version of the maximum entropy theorem is given.

Acknowledgment: The authors would like to express their sincere gratitude to Shun-ichi Amari for his suggestions and the fruitful discussions.

References

[1] Naudts, J. Estimators, escort probabilities, and phi-exponential families in statistical physics. J.Ineq.Pure Appl.Math. 2004, 5, 102.

[2] Amari, S.; Ohara, A.; Matsuzoe, H. Geometry of deformed exponential families: Invariant, dually flat and conformal geometries. Physica A:Statistical Mechanics and its Applications 2012, 391, 4308-4319.

[3] Amari, S.; Ohara, A. Geometry of q-exponential family of probability functions. Entropy 2011, 13, 1170-1185.

[4] Harsha, K.V; Subrahamanian Moosath, K.S. F-geometry and Amari's -geometry on a statistical manifold. Entropy, 2014, 16(5), 2472-2487.

92 Robert Niven, Markus Abel, Steven H. Waldrip, Michael Schlegel and Bernd R. Noack MaxEnt Analysis of Flow and Reaction Networks

The concept of a ``flow network'' -- a set of nodes connected by flow paths -- unites many different disciplines, including electrical, communications, pipe flow, fluid flow, transportation, chemical reaction, ecological and human systems. Historically, the state of a flow network has been analysed by conservation (Kirchhoff's) laws and to some extent by network mappings (e.g.\ Tellegen's theorem), and more recently by numerous dynamical simulation and optimisation methods. A less well explored approach, however, is the use of Jaynes' maximum entropy (MaxEnt) method [1], in which an entropy -- defined over the total uncertainty in the network -- is maximised subject to constraints, to give the stationary state of the network. Several workers have applied MaxEnt methods to the analysis of transportation [2] and hydraulic networks [3], but not always correctly, e.g. in many cases without considering the frictional properties. MaxEnt methods have also been applied directly to network structures (graph ensembles) subject to various configurational constraints (e.g.\ expected degree of each node, k-core structure, etc) [4], but without consideration of flows and potentials on the network.

We present a generalised MaxEnt method to infer the stationary state of a flow network, subject to ``observable'' constraints on expectations of various parameters, as well as ``physical'' constraints arising from conservation laws and frictional properties. The method invokes an entropy defined over all uncertainties within the system, which may include flow rates, potentials, frictional and capacity properties, and the network structure itself. The analysis is quite general, and can incorporate weighted networks (or multiple connections between nodes), sources/sinks at each node, multiple species flows, and network structural constraints. Thus work builds upon a previous MaxEnt analysis of the steady state of infinitesimal flow and dissipative systems [5]. We explore several practical applications, including pipe flow networks, transport networks and chemical reaction networks. This includes the analysis of chaotic, multimolecular chemical reaction networks observed in geological systems (e.g. in ore deposit formation [6]), using a bipartite graph structure. We also show that the latter network formulation provides a natural framework for the analysis of time-varying stationary states as well as steady-state flow.

[1] E.T. Jaynes (G.L. Bretthorst, ed.), Probability Theory:\ The Logic of Science, Cambridge U.P., Cambridge, 2003.

[2] J. de D. Ort\'uzar and L.G. Willumsen, Modelling Transport, 3rd ed., John Wiley \& Sons, Chichester 2001.

[3] T.T. Tanyimboh and A.B. Templeman,

J.\ Op.\ Res.\ Soc.\ 44(4): 383-396 (1993).

[4] J. Park and M.E.J. Newman, Phys. Rev. E 70, 066117 (2004).

[5] R.K. Niven,

Phys.\ Rev.\ E 80(2): 021113 (2009).

[6] A. Ord, B.E. Hobbs and D.R. Lester, Ore Geology Reviews 49, 1-44 (2012).

11:15 – 12:45 Oral session 14 - Bayesian inference

Chaired by: P. Maréchal / Steeve Zozor / J.F. Bercher

14 Bahruz Gadjiev and Tatiana Progulova Origin of generalized entropies and generalized statistical mechanics for superstatistical multifractal systems

So-called generalized entropies are usually introduced in the form of $S_{gen} = \sum \limits _{i=1}^W g(p_i)$ where $g(p_i) \sim \Gamma (1+d, 1-c \ln p_i)$. Then entropies could be qualified by dint of the parameters (c, d) and Boltzmann -- Gibbs entropy $(c = 1, d = 1)$, Shafee entropy $(c = \beta, d = 1)$, Tsallis entropy $(c = 1, d = 0)$ and Anteneodo -- Plastino entropy $\left(c = 1, d = \frac{1}{\eta}\right)$ could be obtained [1].

We consider a multifractal structure as a mixture of fractal substructures and introduce a distribution function of fractal substructure dimensions $f(\alpha)$, where α is a dimension of a fractal substructure. Then we can introduce $g(p) \sim \int \limits _{-\ln p}^{\mu} e^{-y} f(y)dy$ and show that the distribution functions $f(\alpha)$ in the form of $f(\alpha) = \delta (\alpha - 1)$, $f(\alpha) = \delta (\alpha - \theta)$, $f(\alpha) = \frac {1}{\alpha - 1}$, $f(\alpha) = y^{\alpha - 1}$ lead to Boltzmann -- Gibbs, Shafee, Tsallis and Anteneodo -- Plastino entropies conformably. Here $\delta(x)$ is a Dirac delta function. Therefore Shafee entropy corresponds to a fractal structure, Tsallis entropy describes a multifractal structure with a homogeneous distribution of fractal substructures and Anteneodo – Plastino entropy appears in case of a power distribution $f(y)$.

We generalize Fokker -- Planck theory of superstatistics for a multifractal structure. The equation for determination of the distribution function in a fractal substructure has the form:

$\frac{\partial \varphi (x,t)}{\partial t} = \frac{\partial a (x,t) \varphi(x,t)}{\partial x^{\alpha}} - \frac{\partial^2 b(x,t) \varphi(x,t)}{\partial (x^{\alpha})^2}$.

To determine the distribution function in multifractal structures we solve Fokker -- Planck equation

$\frac{\partial \phi (q_1, q_2, t)}{\partial t} = \sum \limits _{k = 1} ^2 \frac {\partial j_k (q_1, q_2, t)}{\partial g_k ^{\alpha}}$

\noindent

where

$j_k(q_1, q_2, t) = K_k (q_1, q_2) - \frac {1}{2} \sum \limits _{l = 1} ^2 Q_{lk} \frac {\partial \phi (q_1, q_2, t)}{\partial g_k ^{\alpha}}$.

Then by the use of Bayes theorem we determine a distribution function for the whole multifractal structure. We compare the results for the distribution functions obtained due to superstatistics approach with the ones obtained according to the principle of the entropy maximum.

We apply the obtained new results to the description of the topology of complex networks.

Reference

1. Hanel R. and Thurner S. When do generalized entropies apply? How phase space volume determines entropy, Europhysics Letters 96, 50003 (2011)

40 Hiroshi Matsuzoe Information geometry of Bayesian statistics

Information geometry is a differential geometric approach of statistical inferences. The study of information geometry was originated of the results by Hotelling (1929) and Rao (1945) in having considered a statistical model to be a Riemann manifold. After a formulation of dual affine connections on statistical models by Amari and Nagaoka (1982), information geometry has been applied various fields of mathematical sciences. In particular, a Riemannian manifold with mutually dual flat affine connection is called a dually flat space. This geometric structure has close relations among geometry of statistical inferences and relative entropies.

Information geometry of Bayesian statistics has been studied by Komaki (1996), Takeuchi and Amari (2005), and Takeuchi, Amari and Matsuzoe (2006), etc. In Bayesian statistics, a prior distribution on a parameter space plays an important role. From a viewpoint of information geometry, a prior distribution is regarded as a volume element on a manifold of statistical model since a parameter space is a local coordinate system of a statistical model. For the arguments of projected Bayesian estimators and bias correction of estimators, differential geometric gradient vector fields of volume elements, called the Tchebychev vector fields, are important.

On the other hand, a method of statistical inference based on non-additive entropy has been studied recently, which is called the Tsallis statistics. In this method, deformed probability distributions, called escort distributions, play important roles. Dually flat structures on deformed exponential families, generalized maximum likelihood methods, and generalized relative entropies are introduced from escort distributions. Deformations of probability measures on sample spaces are important in Tsallis statistics while deformations of measures on parameter spaces are important in Bayesian statistics.

In this presentation, we give a survey about information geometry of Bayesian statistics. Then we consider relations between Bayesian statistics and Tsallis statistics.

103 R. Preuss and Udo V. Toussaint Bayesian uncertainty quantification for an electrostatic plasma model

The numerical simulation of fusion plasmas involves field quantities which are hampered by noise, some of which in time and/or space. In order to compare results from experiment and model, or to have an estimation of the fluctuation margin of a model prediction, an uncertainty quantification is necessary. We employ nonintrusive *polynomial chaos expansion to generate a functional representation which allows to investigate the uncertainty propagation through the forward model. An instructive example of absorption in media serves for the validation of the procedure. Finally the method is applied to the Vlasov-Poisson model describing electrostatic plasmas in one and two dimensions.*

14:00 – 17:00 Oral session 15 - Maximum Entropy Principle

Chaired by: John Skilling / Hiroshi Matsuzoe

3 Alexander Fradkov and Dmitry Shalymov Dynamics of non-stationary processes that follow the MaxEnt principle for Shannon and Renyi entropies

Although information theory was initially based on some concepts of statistical physics, now, thanks to Claude Shannon, the information approach can be taken as a basis to develop statistical physics and thermodynamics. In 1948, Shannon introduced his information entropy for an absolutely continuous random variable x having probability density function (pdf) p. In 1961, Alfred Renyi generalized Shannon entropy as a one parameter family of entropies. Renyi entropy is widely used nowadays in communication and coding theory, signal processing, data mining and many other areas.

A phenomenon when system tends to its state of maximum entropy is known as the maximum entropy (MaxEnt) principle. Since seminal works of E.T. Jaynes (1957) and until recent years the MaxEnt principle attracts a strong interest of researchers. The MaxEnt principle defines the asymptotic behavior of the system, but does not say anything about how the system moves to an asymptotic behavior. Despite a large number of publications studying the maximum entropy states, the dynamics of evolution and transient behavior of the systems are still not well investigated.

In this paper we have considered dynamics of pdf for non-stationary processes that follow MaxEnt principle. We have derived a set of equations describing dynamics of non-stationary (transient) states and describing a way and trajectory of the system that tends to the state with maximum entropy. Systems with discrete probability distribution and continuous pdfs are considered under mass conservation and energy conservation constraints. The uniqueness of the limit pdf and asymptotic convergence of pdf are examined. We have proposed a generalized form of Renyi distribution (RD) for an arbitrary number of constraints and have shown that this distribution corresponds to the state with maximum Renyi entropy.

Convergence of pdfs does not lead to the convergence of the corresponding differential entropies. Based on sufficient conditions the nontrivial convergence of differential entropies has also been proved.

We use the speed-gradient (SG) principle originated in control theory [1]. Applicability of the SG principle has already been experimentally tested for the systems of finite number of particles simulated with the molecular dynamics method [2,3]. We apply similar approach for the systems with discrete and continuous probability distributions to maximize Shannon and Renyi entropies.

The equations derived on the basis of the SG principle allow forecasting the dynamics of non-stationary systems. If a goal function is the entropy of a system then the extreme

SG principle supplements Gibbs and Jaynes MaxEnt principle to determine the direction of evolution of the system when it tends to the state with maximum entropy.

[l] A. Fradkov, Physics-Uspekhi 48, 103-127 (2005).

[2] A. Fradkov, Entropy 10, 757-764 (2008).

[3] A. Fradkov, A. Krivtsov, MaxEnt 2010 1305, 399-406 (2010).

16 Patrick Bogaert and Sarah Gengler MinNorm approximation of MaxEnt/MinDiv problems for probability tables

Categorical data are found in a wide variety of important applications in environmental sciences and dealing with multivariate analyses is a challenging topic. Rebuilding a multivariate probability table becomes an issue and is expected to lead to poor probability estimates when a very limited number of samples are at hand. In order to take into account the lack of data, the information can be rewritten as inequality constraints instead of using the few sampled values as direct probability estimates. There is thus a need for an efficient method that allows us to rebuild a multivariate probability table from equalities and inequalities constraints. The methodology suggested by Bogaert (2011) and presented in this paper is an extension of Bayesian Data Fusion (BDF) (Bogaert and Fasbender, 2007) and the Bayesian Maximum Entropy (BME) (Christakos, 2002).

Rebuilding a probability function from equalities constraints can be done through a classical MaxEnt methodology. Bogaert (2011) shows how MaxEnt problem can be implemented by using iterated minimum norm (MinNorm) approximations. Minimum divergence (MinDiv) methodology extends the problem to the case of inequalities constraints and, again, Bogaert (2011) shows that MinNorm approximations can be applied and iterated. Thus, iterated MinNorm approximations are a fast and efficient way to combine equalities and inequalities constraints to rebuild a multivariate probability table (Bogaert, 2011).

MinNorm methodology for solving problems involving both equalities and inequalities constraints can be applied in a wide variety of applications. MinNorm approximations become useful, for instance, when only few data are available or when taking into account experts opinion rewritten as equalities and inequalities constraints is of prime interest in probability estimates. Few simple examples are presented in order to illustrate the benefits of the methodology (Wahyudi et al., 2013).

References :

Bogaert, P, 2011. Fast implementationof the MaxEnt/MinDiv Problem for Probability Tables using Iterated MinNorm Approximations and Topological Properties. Working Paper, Université catholique de Louvain.

Bogaert, P., Fasnebder, D., 2007. Bayesian Data Fusion in a Spatial Prediction Context: A General Formulation. Stochstic Environmental Research and Risk Assessment, vol. 21. Springer, pp.695-709.

Christakos, G., Bogaert, P., Serre, M., 2002. Temporal GIS: Advanced Functions for Field-Based Applications. Springer.

Wahyudi, A., Bartzke, M., Küster, E., and Bogaert, P. (2013). Maximum entropy estimation of a benzene contamined plume using ecotoxological assays. Environmental Pollution, 172:170-179.

42 Shinto Eguchi, Osamu Komori and Atsumi Ohara Duality in a maximum generalized entropy model

We discuss a special class of generalized entropy by the use of generator functions. The generator function U(s) defined on the real axis is assumed to be convex and strictly increasing, which leads to the conjugate convex function. The generalized entropy for a probability density function is defined by U. Let us fix a canonical variable t(x) of d-dimension. Then we consider a mean equal space of all probability density functions such that the expectation of t(x) is constant vector. The maximum generalized entropy distribution under the constraint to the mean equal space is given by a simple form. In fact, if U is an exponential function, then the entropy reduces to the Boltzmann-Shannon entropy, and the maximum entropy distribution belongs to an exponential family. We discuss the differential geometric property for the model. For this we introduce the Riemannian metric and dual linear connections induced from the divergence associated with the generalized entropy. Then we will show that the model of maximum entropy distributions is totally geodesic, that is, for any pair of distributions of the model the geodesic curve connecting the two distributions with respect to either of the dual connections is again included in the model. Furthermore, we consider a special property for two parameters defined by the mean equality. In effect the natural parameter and expectation parameter are affine parameters with respect the dual connections in which the two parameters associates with two convex functions with conjugate convexity. We will show that the Riemannian metric is expressed by the Hessian matrices for the two convex functions with respect to the two parameters. We observe that this duality between the two parameters is derived from the conjugacy between generator functions U(s) and the conjugate function.

58 Diego González, Sergio Davis and Gonzalo Gutiérrez Newtonian Dynamics from the principle of Maximum Caliber

In this work we show that the non-equilibrium generalization of Jaynes' principle of Maximum Entropy, the Principle of Maximum Caliber, when applied to the unknown trajectory followed by a particle, leads to Newton's second law under two quite intuitive assumptions (the expected square displacement in one step and the spatial probability distribution of the particle are known at all times). Our derivation explicitly highlights the role of mass as an emergent measure of the fluctuations in velocity (inertia) and the origin of potential energy as a manifestation of spatial correlations. Our findings provide support to the application of Newton's equations outside mechanical (or even physical) systems, such as modelling ecological, financial and biological systems.

NOTES

NOTES

MAXENT'2014 PROGRAM AT A GLANCE

	Sunday, September 21	Monday, September 22	Tuesday, September 23	Wednesday, September 24	Thursday, September 25	Friday, September 26
MORNING	REGISTRATION 9:00 - 10:00	REGISTRATION 8:00 - 9:00				
	TUTORIAL 1 10:00 - 13:00	**OPENING & SESSION 1** **9:15 - 11:15**	**SESSION 4** **9:00 - 10:45**	**SESSION 7** **9:00 - 10:15**	**SESSION 10** **9:00 - 10:30**	**SESSION 13** **9:00 - 11:00**
		Coffee break 11:15 - 11:30	Coffee break 10:45 - 11:00	Coffee break 10:15 - 10:30	Coffee break 10:30 - 10:45	Coffee break 11:00 - 11:15
		POSTER SESSION 1 **11:30 - 12:45**	**SESSION 5** **11:00 - 12:30**	**SESSION 8** **10:30 - 12:45**	**SESSION 11** **10:45 - 12:45**	**SESSION 14** **11:15 - 12:45**
LUNCH BREAK	Lunch break 13:00 - 14:30	Lunch break 12:45 - 14:00	Lunch break 12:30 - 14:00	Lunch break 12:45 - 14:00	Lunch break 12:45 - 14:00	Lunch break 12:45 - 14:00
AFTERNOON	TUTORIAL 2 14:30 - 19:30	**SESSION 2** **14:00 - 15:00**	**SESSION 6** **14:00 - 15:30**	**SESSION 9** **14:00 - 15:00**	**SESSION 12** **14:00 - 15:30**	**SESSION 115** **14:00 - 17:00**
		Coffee break 15:00 - 15:30	Coffee break 15:30 - 15:45	SOCIAL ACTIVITIES from 16:00	Coffee break 15:30 - 15:45	Coffee break 17:00 - 17:15
		SESSION 3 **15:30 - 17:30**	**POSTER SESSION 2** **15:45 - 18:45**		**POSTER SESSION 2** **15:45 - 18:45**	
EVENING	WELCOME COCKTAIL at the CLOS LUCE 19:30 - 21:30	WELCOME COCKTAIL at the CLOS LUCE 18:00 - 19:30		GALA DINNER AT CHÂTEAU DE CHISSAY		

BOISE STATE UNIVERSITY

entropy

Société Mathématique de France

SMF

SMAI

CNRS

Supélec

see

UNIVERSITÉ PARIS SUD

Comprendre le monde, construire l'avenir®