ihp Institut Henri Poincaré

scai SORBONNE CENTER FOR ARTIFICIAL INTELLIGENCE

see since 1883

cnrs

International Conference on Bayesian and Maximum Entropy methods in Science and Engineering

41st MaxEnt2022 Conference

JULY 18-22, IHP, PARIS

PROGRAM

SPONSORS & PARTNERS

41st MaxEnt2022 Conference

JULY 18-22, IHP, PARIS

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 2022 conference. The forty-first 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)rganized by SEE (https://www.see.asso.fr/en) and SCAI Sorbonne University at Institut Henri Poincaré, in Paris from July 18th to 22nd, 2022. Proceedings of MaxEnt 2022 Proceedings will be published by MDPI.

MDPI Entropy publication is an editorial partner to publish long papers issue from MaxEnt 2022 Tutorial, Invited and selected authors in special issue. All other interested authors are invited to submit their long papers in this Special Issue:

https://www.mdpi.com/journal/entropy/special_issues/MaxEnt2022

MaxEnt 2022 will take place in Institut Henri Poincaré (http://www.ihp.fr/en). Since its creation in 1928, the IHP has taken an interest in all the disciplinary fields, especially at the interface between mathematics and theoretical physics. The Henri Poincaré Institute was inaugurated on November 17, 1928. This institute is organized around two missions: a teaching mission with the Chairs of Probability and Mathematical Physics and Physical Theory and a research mission with the invitation French and foreign scientists to give lectures in the field of physics and mathematics. Thanks to these two missions, a small group of mathematicians (Emile Borel, Maurice Fréchet, Georges Darmois) uses this institute to create a scientific field dedicated to the theory of probability in Paris. At the same time, they are using this institute to acquire a remarkable place on the international probabilistic scene. First lectures at IHP were given by Emile Borel, Maurice Fréchet and Léon Brillouin on Probability.

MaxEnt 2022 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 on Geometric Structures of Heat, Information and Entropy.

This year special interest will be more focused on:

- **Foundations of probability, inference, information, entropy**
- **Bayesian Physics-Informed & Thermodynamics-Informed Machine Learning**
- **Information theory, machine learning tools for inverse problems**
- **Bayesian and Maximum Entropy in real world applications**
- **Geometric Statistical Mechanics/Physics, Lie Groups Thermodynamics & Maximum Entropy Densities**
- **Quantum: Theory, Computation, Tomography and Applications**

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 Monday Tutorials:

› **Ali Mohammad-Djafari** (CNRS, France) - *Bayesian and Machine Learning Methods for Inverse Problem*

› **Kevin H. Knuth** (University at Albany, USA) - *Why Mathematics Works and Why Physics is Mathematical*

› **John Skilling** (University of Cambridge, UK) - *Foundations*

› **Frank Nielsen** (Sony CSL, Japan) - *Introduction to Information Geometry*

› **Fréderic Barbaresco** (THALES, France) - *Symplectic Theory of Heat and Information based on Souriau Lie Groups Thermodynamics, Coquinot Thermodynamic Dissipative Bracket and Sabourin Transverse Poisson Structures: Applications to Lindblad Equation*

› **Ariel Caticha** (University at Albany, USA) - *Entropic Dynamics and Quantum Measurement*

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

› **Anna Simoni** (ENSAE, France) - *Bayesian Exponentially Tilted Empirical Likelihood to Endogeneity Testing*

› **Antoine Bourget** (CEA and ENS Paris, France) - *The Geometry of Quivers*

› **Bobak Toussi Kiani** (MIT, USA) - *Quantum algorithms for group convolution, cross-correlation, and equivariant transformations*

› **Emtiyaz Khan** (RIKEN, Japan) - *The Bayesian Learning Rule*

› **Fabrizia Guglielmetti** (ALMA Regional Center Scientist at European Southern Observatory, Germany) - *Bayesian and Machine Learning methods in the Big Data era for astronomical imaging*

› **Livia Partay** (University of Warwick, UK) - *Nested sampling for materials*

› **Lorenzo Valzania** (LKB: Sorbonne University - ENS - Collège de France, France) - *Imaging behind scattering layers*

› **Pierre-Henri Wuillemin** (Laboratoire d'Informatique de Paris, France) - *Learning Continuous High-Dimensional Models using Mutual Information and Copula Bayesian Networks*

› **Torsten Ensslin** (MPA, Germany) - *Theoretical Modeling of Communication Dynamics*

› **Will Handley** (University of Cambridge, UK) - *Bayesian sparse reconstruction: a brute-force approach to astronomical imaging and machine learning*

› **Piotr Graczyk** (Angers, France) - *Graphical Gaussian models associated to a homogeneous graph with permutation symmetries*

› **Olivier Rioul** (Telecom ParisTech) - *What is Randomness? The Interplay between Alpha Entropies, Total Variation and Guessing*

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.

MAXENT 2022 COMMITTEES

General chairs of MaxEnt 2022

Frédéric Barbaresco
SEE ISIC Club President
THALES, France

Ali Mohammad-Djafari
Former Research Director of CNRS,
Founder and Director of «International Science Consulting and Training (ISCT)», France

Frank Nielsen
Sony Computer Science Laboratories Inc
Tokyo, Japan

Martino Trassinelli
INSP (Institut des NanoSciences de Paris)
CNRS Sorbonne Université, France

Local organizing committee

- **Xavier Fresquet,** Sorbonne Center for AI (SCAI) Deputy Director
- **Lune Maillard,** INSP (Institut des NanoSciences de Paris)
- **Maria Trocan,** Institut Supérieur d'Électronique de Paris
- **Pierre-Henri Wuillemin,** LIP6 UMR 7606 Sorbonne Université - CNRS

Secretary

› **Marianne Emorine,** SEE, France

Scientific committee

› **Romke Bontekoe,** Bontekoe Research, The Netherlands
› **Ariel Caticha,** University at Albany, USA
› **Keith Earle,** University at Albany, USA
› **Paul Goggans,** University of Mississippi, USA
› **Fabrizia Guglielmetti,** European Southern Observatory, Germany
› **Kevin Knuth,** University at Albany, USA
› **Donald Nelson,** USA
› **Robert Niven,** The University of New South Wales, Australia
› **Arthur Parzygnat,** Institut des Hautes Études Scientifiques, France
› **John Skilling,** University of Cambridge, UK
› **Julio Michael Stern,** University of São Paulo, Brazil
› **Barrie Stokes,** Australia
› **Geert Verdoolaege,** University of Ghent, Belgium
› **Wolfgang von der Linden,** Graz University of Technology, Austria
› **Udo von Toussaint,** Max-Planck-Institute for Plasma Physics, Germany

ABOUT SEE

see
since 1883

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.

MAXENT 2022 TUTORIAL SPEAKERS

Ali Mohammad-Djafari
L2S, CNRS, CentraleSupélec-Univ Paris Saclay, France
International Science Consulting and Training (ISCT), France
Scientific leader of Shanfeng company, Shaoxing, China

Bayesian Inference, Machine Learning and Physics-Informed Neural Networks methods for Inverse Problems

Abstract: Inverse problems arise anywhere we have indirect measurement. Classical methods for them are mainly based on regularization theory, in particular those which are based on optimization of a criterion with two parts: a data-model matching criterion and a regularization term. Different criteria for these two terms and a great number of standard and advanced optimization algorithms have been proposed and used with great success. When these two terms are distances or divergence measures, they can have a Bayesian Maximum A Posteriori (MAP) interpretation where these two terms correspond, respectively, to the likelihood and prior probability models.

The Bayesian approach gives more flexibility in choosing these terms via the likelihood and the prior probability distributions. This flexibility goes much farther with the hierarchical models and appropriate hidden variables. Also, the possibility of estimating the hyperparameters gives much more flexibility for semi-supervised methods.

However, the full Bayesian computations can become very heavy computationally, in particular when the forward model is complex and the evaluation of the likelihood needs high computational cost. Using surrogate simpler models can become very helpful to reduce the computational costs, but then, we have to account for uncertainty quatification (UQ) of the obtained results.

The Machine Learning (ML) methods and algorithms have gained great success in many image processing tasks, such as classification, clustering, object detection, semantic segmentation, etc. These methods are mainly based on Neural Networks (NN) and in particular Convolutional NN (CNN), Deep NN, etc. Using these methods directly for inverse problems, as intermediate pre-processing or as tools for doing fast approximate computation in different steps of regularization or Bayesian inference have also got success, but not as much as they could.

Recently, the Physics-Informed Neural Networks have gained great success in many inverse problems, proposing interaction between the Bayesian formulation of forward models, optimization algorithms and ML specific algorithms for intermediate hidden variables. These

methods have become very helpful to obtain approximate practical solutions to inverse problems in real world applications.

In this tutorial paper, first, particular examples of simple image denoising, more difficult image restoration and Computed Tomography (CT) image reconstruction will illustrate this cooperation between ML and Inversion methods. Then, some examples on more complex non linear imaging systems opens the perspectives on these methods.

Kevin H. Knuth
University at Albany, USA

Why Mathematics Applies to the Real World and a New Approach to Foundations

Abstract: Many of us have a tendency to think of the laws of physics as mathematical laws dictated by Nature. But this raises serious questions as to why these laws and not others? How are such laws enforced and why can't they be broken? Furthermore, given that mathematics is a human invention, why should these laws be mathematical in nature?

We demonstrate that some of the laws of physics arise from the fact that we find utility in describing things by mapping attributes to numbers. Careful examination reveals that algebraic symmetries constrain the ways in which such numbers can be consistently assigned resulting in constraint equations. These constraint equations are what we generally think of as the physical laws, while it is the underlying symmetries that form the foundation. The laws ensure that the mathematics does not violate the foundational symmetries. As long as those symmetries faithfully describe properties of the real world, our mathematics is guaranteed to work by design!

More importantly, this realization results in a new, unusual, and unexpected approach to the foundations of physics where one derives the subset of mathematically possible laws (constraint equations) given known or hypothesized symmetries. Potential advantages to this new approach are discussed here; whereas several surprising results are discussed in the following tutorial by John Skilling.

John Skilling
University of Cambridge, UK

Foundations

Abstract: As physicists, we wish to make mental models of the world around us. For this to be useful, we need to be able to classify features of the world into symbols and develop a rational calculus for their manipulation. In seeking maximal generality, we aim for minimal restrictive assumptions. That inquiry starts with basic arithmetic and proceeds to develop without further ado the formalism of quantum theory and relativity.

Frank Nielsen
Sony CSL, Japan

Introduction to Information Geometry

Abstract: Information geometry studies geometric structures, statistical distances, and the concepts of statistical invariance of a family of probability distributions called the statistical model.

A statistical parametric model can be handled as a Riemannian manifold equipped with the Fisher metric inducing the Rao distance.

This Fisher-Rao Riemannian structure can further be generalized into dualistic structures relying on a pair of torsion-free affine connections coupled to the Fisher metric.

These dual structures explain the relationships between statistical estimators like the maximum likelihood estiomator and statistical models like exponential families which can be obtained from the maximum entropy principle.

In particular, we shall highlight the dually flat structures with their generalized Pythagoras theorems allowing one to interpret divergence minimizers as information projections.

Several applications of information geometry in information sciences and machine learning will be described.

Fréderic Barbaresco
THALES, France

Symplectic Theory of Heat and Information based on Souriau Lie Groups Thermodynamics, Coquinot Thermodynamic Dissipative Bracket and Sabourin Transverse Poisson Structures: Applications to Lindblad Equation

Abstract: Jean-Marie Souriau has extended the classical notion of Gibbs' canonical ensemble to the case of a symplectic manifold on which a Lie group has a Hamiltonian action. Souriau model is a new symplectic theory of heat and of Information Geometry, called "Lie Groups Thermodynamics". This model gives an archetypal, and purely geometric, characterization of Entropy, which appears as an invariant Casimir function in coadjoint representation, from which we will deduce a geometric heat equation.

The approach also allows generalizing the Fisher metric of information geometry thanks to the 2-form KKS (Kirillov, Kostant, Souriau) in the affine case via the Souriau's cocyle. The dual space of the Lie algebra foliates into coadjoint orbits that are also the level sets on the entropy. This Lie Groups Thermodynamics could be interpreted in the framework of Thermodynamics by the fact that motion remaining on these symplectic leaves is non-dissipative, whereas motion transversal to these symplectic leaves is dissipative. In a second part, we will address dissipative symplectic theory of Heat and Information by considering the transverse Poisson structure to the symplectic leaf of the dual space, endowed with its canonical Poisson structure, based on the metriplectic bracket ensuring both conservation of energy and non-decrease of entropy. More recently, Baptiste Coquinot has introduced a general point of view from non-equilibrium thermodynamics to derive new foundation theory of dissipative brackets, considering a set of variables that appear more natural for constructing the metriplectic bracket and then presenting a systematic way to derive general dissipative brackets. Baptiste Coquinot's construction shows that

the dissipative brackets are completely natural for non-equilibrium thermodynamics, just as Poisson brackets are natural for Hamiltonian dynamics, deriving a general dissipative bracket, for the first time, from basic thermodynamic first principles. We will explore links to transverse structure from Dirac's theory of dissipative brackets for constrained hamiltonian systems. We will explore another approach than Metriplectic approach, based on the transverse Poisson structure that has been studied by Michel Saint-Germain in his PhD inspired by Fokko du Cloux seminal works. In continuity, Hervé Sabourin has studied polynomial nature of Poisson structures transverse to nilpotent adjoint orbits and has proved that restriction to the transverse slice of the Casimir functions of the Lie-Poisson structure are independent Casimir functions of the transverse Poisson structure, and has also proved that there are two polynomial Poisson structures on the transverse slice to the symplectic leaf, which have Casimir functions, namely the transverse Poisson structure and the determi-nantal structure, constructed by using these Casimirs. To illustrate previous models, we will consider for application the Lindblad equation describing dissipative versions of the Hamiltonian Liouville equation of the density matrix, using that the dissipative part of Lindblab operator has been described as a gradient of the relative entropy, that equilibria can be obtained by the maximum entropy principle.

Ariel Caticha
University at Albany, USA

Entropic Dynamics and Quantum Measurement

Abstract: The problem of quantum measurement is intimately associated with most of those features of quantum theory that make it so strange and fascinating. Does the quantum state reflect incomplete information or is it something real, ontic? If the latter, can wave functions undergo a physical process of collapse during measurement? Alternatively, if no collapses ever occur, and wave functions always obey the linear Schrödinger equation, how do quantum measurements ever yield definite outcomes? How does one negotiate the interface between the microscopic quantum world and the macroscopic classical world of the measuring device? Finally, do at least some privileged variables represent something real with definite values at all times? Or, alternatively, are the values of all observables created during the act of measurement? If so, how can one ever say that anything real exists when nobody is looking?

Entropic Dynamics (ED) is an approach to quantum mechanics that is ideally suited to tackle the questions above because it is based on entropic and Bayesian methods of inference that have been designed to process information and data. After a brief overview of ED we discuss the direct measurement of microscopic positions including their amplification to achieve observability at the macroscopic level. Then we discuss the "indirect measurement" of other observables that require more elaborate devices including the classical von Neumann measurements and also the so-called weak measurements.

The success of the ED approach hinges on a clear ontological commitment: In ED the positions of particles enjoy the privileged role of being the only ontic variables while probabilities and wave functions are fully epistemic in nature. Thus, positions reflect real properties with definite values that are not created by the act of measurement. In contrast, all other observables are epistemic because they reflect properties of the wave function. This explains how it is that their values are "created" by the act of measurement.

History of probability

Laurent Mazliak
Sorbonne Université, LPSM

Borel and the emergence of probability on the mathematical scene in France

Abstract: In 1928 the Henri Poincaré Institute opened in Paris thanks to the efforts of the mathematician Emile Borel and the support of the Rockefeller Foundation. The teaching and research on the mathematics of chance were placed by Borel at the center of the institute's activity, a result imposed by the French mathematicians in the face of indifference and even hostility towards a discipline accused of lack of seriousness. This historical account, based in large part on the results of Matthias Cléry's thesis, presents the way in which Borel became convinced of the importance of making up for the gap between France and other countries as regards the place of probability and statistics in the educational system, and elaborated the strategy that led to the creation of the IHP and how its voluntarist functioning enabled it to become in ten years one of the main world centres of reflection on this subject.

SOCIAL EVENTS

Gala Dinner at the Procope restaurant

The Gala Dinner will take place on Tuesday July 19th, 2022 from 19:00.

Address: 13 Rue de l'Ancienne Comédie, 75006 Paris
12 minutes' walk from IHP
Metro station: Odéon

Le Procope, the oldest café in Paris in the heart of Saint-Germain des Prés.

A mythical restaurant in the 6th arrondissement of Paris since 1686, Le Procope is a place steeped in history where the greatest writers and intellectuals met (Rousseau, Diderot, Verlaine...) with a historic decor.

The gala dinner is included in the full registration fees only. Additional tickets are available for sale for accompanying persons (ticket price: € 80,-).

GUIDED TOURS

Sightseeing tours will take place on Thursday July 21st, 2022 from 18:00 to 19:30 (duration 1h30). For registered participants only.

TOUR 1: Ile de la Cité

Meeting point: Notre-Dame cathedral square at 18:15 (17 min walk from IHP)

The "île de la cité" is the historical cradle of France's capital city. Once populated by the Celtic tribe of the Parisii, this urban enclave presents a double perspective. On the west side, the tip of the island "la Pointe du vert-galant" and the Place Dauphine with its "provincial charm". But also the Palace of Justice sheltering the Sainte-Chapelle. Passing through the flower market, the visitor will then discover the east side of the island with the famous Notre-Dame Cathedral and the few medieval lanes that survived the major urban changes of Paris in the nineteen's century, to enjo y a stroll in the heart of the ancient city.

Extra tickets are available for sale at the Registration Desk (ticket price: € 10,-).

TOUR 2: Rue de la Montagne Sainte-Geneviève

Meeting point: Place du Panthéon at 18:00 (550m from IHP)

Back in Medieval Times, the Montagne Sainte-Geneviève sheltered the first Parisian schools and universities. The intellectual vocation for transmission that was born on that hill thrived with time. It was then consecrated during the Revolution and the following period of the 1st Empire with the opening of the Ecole Polytechnique in 1804, the Ecole Normale Supérieure of the « rue d'Ulm » in 1808. Then came the Campus Pierre et Marie Curie, the Sainte-Geneviève Library, the Lycée Henri IV, the Panthéon… This walking tour will guide you from one to another of those many prestigious Parisian institutions dedicated to knowledge.

Extra tickets are available for sale at the Registration Desk (ticket price: € 10,-).

MAXENT 2022 PROGRAM

Monday, July 18th

08.30-09.00	**Registration**
09.00-09.20	**Tutorial 01:** Ali M. Djafari
09.20-09.40	
09.40-10.00	
10.00-10.20	**Tutorial 02:** Kevin H. Knuth
10.20-10.40	
10.40-11.00	
11.00-11.20	**Coffee Break**
11.20-11.40	**Tutorial 03:** John Skilling
11.40-12.00	
12.00-12.20	
12.20-12.40	**Lunch Break**
12.40-13.00	
13.00-13.20	
13.20-13.40	**Tutorial 04:** Frank Nielsen
13.40-14.00	
14.00-14.20	
14.20-14.40	**Tutorial 05:** Frederic Barbaresco
14.40-15.00	
15.00-15.20	
15.20-15.40	**Coffee Break**
15.40-16.00	**Tutorial 06:** Ariel Caticha
16.00-16.20	
16.20-16.40	
16.40-17.00	**Coffee Break**
17.00-17.20	**History of probability** Laurent Mazliak
17.20-17.40	
17.40-18.00	
18.00-18.20	**Welcome Cocktail**
18.20-18.40	
18.40-19.00	

Tuesday, July 19th

08.30-09.00	**Registration**
09.00-09.20	**Invited Speaker 01** Emtiyaz Khan (87)
09.20-09.40	
09.40-10.00	**71** - Fábio C.C. Meneghetti
10.00-10.20	**80** - Matteo Guardiani
10.20-10.40	**75** - Keiko Uohashi
10.40-11.00	**Coffee Break**
11.00-11.20	**Invited Speaker 02** Torsten Ensslin (23)
11.20-11.40	
11.40-12.00	**86** - Christina Pawlowitsch
12.00-12.20	**32** - Andrew Charman
12.20-12.40	**Lunch Break**
12.40-13.00	
13.00-13.20	
13.20-13.40	**Invited Speaker 03** Lorenzo Valzania (52)
13.40-14.00	
14.00-14.20	**31** - Berlin Chen
14.20-14.40	**79** - Yannis Kalaidzidis
14.40-15.00	**74** - Geoffroy Delamare

15.00-15.20	**Coffee Break**
15.20-15.40	**38** - Florent Leclercq
15.40-16.00	**70** - Orestis Loukas
16.00-16.20	**67** - Łukasz Tychoniec
16.20-16.40	**17** - Riko Ketler
16.40-17.00	**Coffee Break**
17.00-17.20	**19** - Philipp Joppich
17.20-17.40	**27** - Philippe Jacquet
17.40-18.00	**82** - Mariela Portesi
18.00-18.20	**34** - François Verdeil
18.20-18.40	**Gala Diner** Le Procope
18.40-19.00	
19.00-22.00	

Legend:

Topic 3 Information theory, machine learning tools for inverse problems

Topic 4 Bayesian and Maximum Entropy in real world applications

Topic 6 Quantum: Theory, Computation, Tomography and Applications

08:30 – 09:00 **Registration**

09:00 – 10:40

Session 1: Information Theory, Information Geometry, And Machine Learning

Chaired by: John Skilling / Ali Mohammad-Djafari

87 - Invited Speaker 01: Emtiyaz Khan - *The Bayesian Learning Rule*

Abstract: In this talk, I will show that solutions of Bayesian problems have deep roots in information geometry. This perspective is then used to show that a wide-variety of machine-learning algorithms are instances of a single learning-rule called the Bayesian learning rule. The key idea in deriving such algorithms is to approximate the posterior using candidate distributions estimated by using natural gradients. Different candidate distributions result in different algorithms and further approximations to natural gradients give rise to variants of those algorithms. The new perspective can be used to design a new ''adaptive" learning-paradigm that could solve many challenges faced by deep-learning researchers. The talk is based on Khan and Rue (2022).

71 - Fábio C. C. Meneghetti, Henrique K. Miyamoto and Sueli I. R. Costa - *Information Properties of a Random Variable Decomposition through Lattices*

Abstract: A full-rank lattice in the Euclidean space is a discrete set given by all integer linear combinations of a basis. Given a continuous probability distribution, two operations can be induced by considering the quotient of the space by such lattice: wrapping and quantization. The wrapped distribution over the quotient is obtained by summing the density over each coset, while the quantized distribution over the lattice is defined by integrating over each fundamental domain translation. The former is used in the definition of the flatness factor, which is, up to a constant, the $L\infty$ distance from a wrapped probability distribution to a uniform one, and is important for the analysis of lattice-based cryptography and of lattice coset coding, particularly for the AWGN and wiretap channels. The latter is related to closest-vector decoding, for special fundamental domains. Given a continuous random variable on , a lattice Λ, and a fundamental domain D which tiles through Λ, we consider the induced wrapped and quantized random variables over D and Λ respectively, and show that they sum up to the original one. We investigate information properties of this decomposition, such as entropy, mutual information and the Fisher information matrix, and show that this approach can be extended to the more abstract context of locally compact topological groups, in particular to Lie groups, where the concepts of lattice and probability density can be naturally defined.

80 - Andrija Kostić, Philipp Frank, Matteo Guardiani, Sebastian Hutschenreuter, Maximilian Kurthen, Reimar Leike and Torsten Enßlin - *Towards Moment-constrained Causal Modeling*

Abstract: The fundamental problem of causal inference is to discover causal relations between variables used to describe observational data. We address this problem within

the formalism of Information Field Theory (IFT). Specifically, we focus on the problems of bivariate causal discovery ($X \to Y$, $Y \to X$) from an observational dataset (X, Y). The bivariate case is especially interesting because the methods of statistical-independence testing are not applicable here. For this class of problems, we propose the Moment-constrained Causal Model (MCM). The MCM goes beyond the Additive Noise Model by exploiting Bayesian hierarchical modeling to provide non-parametric reconstructions of the observational distributions. In order to identify the correct causal direction, we compare the performance of our newly-developed Bayesian inference algorithm for different causal directions ($X \to Y$, $Y \to X$) by calculating the evidence lower bound (ELBO). To this end, we develop a new method for the ELBO estimation that takes advantage from the adopted variational inference scheme for parameter inference.

75 - Keiko Uohashi - *A foliation of probability simplexes for transition of α-parameters*

Abstract: This paper treats dualistic structures of probability simplexes from the point view of information geometry. We investigate a foliation of probability simplexes for transition of α-parameters, not for a fixed \alpha-parameter. Properties of divergences on the foliation are also described when different α-parameters are defined on different leaves.

10:40 – 11:00 **Coffee break**

11:00 – 12:20

Session 2: Inference And Information In Human Activities

Chaired by: Sacha Ranftl / Ali Mohammad-Djafari

23 - Invited Speaker 02: Torsten Ensslin - *Reputation communication from an information perspective*

Abstract: Communication, the exchange of information between intelligent agents, be they equipped with human or artificial intelligence, is susceptible to deception and misinformation. Reputation systems can help agents decide to what extent to trust an information source that is not necessarily reliable. Consequently, the reputation of the agents themselves determines the influence of their communication on the beliefs of others. This makes reputation a valuable resource and thus a natural target for manipulation. To study the vulnerability of reputation systems, we simulate the dynamics of communicating agents seeking high reputation within their social group using an agent-based model. The simulated agents are equipped with a cognitive model that is bounded in mental capacity but otherwise follows information-theoretic principles. Various malicious strategies of agents are examined for their effects on group sociology, such as sycophancy, egocentricity, pathological lying, and aggressiveness. Phenomena resembling real social psychological effects are observed, such as echo chambers, self-deception, deceptive symbiosis, narcissistic supply, and freezing of group opinions.

86 – Christina Pawlowitsch – *Strategic Manipulation in Bayesian Dialogues*

Abstract: In a Bayesian dialogue two individuals report their Bayesian updated belief about a certain event back and forth, at each step taking into account the additional information contained in the updated belief announced by the other at the previous step. Such a process, which operates through a reduction of the set of possible states of the world, converges to a commonly known posterior belief, which can be interpreted as a dynamic foundation for Aumann's agreement result. Certainly, if two individuals have diverging interests, truthfully reporting one's Bayesian updated belief at every step might not be optimal. This observation could lead to the intuition that always truthfully reporting one's Bayesian updated belief were the best that two individuals could do if they had perfectly coinciding interests and these were in line with coming to know the truth. This article provides an example which shows this intuition to be wrong. In this example, at some step of the process, one individual has an incentive to deviate from truthfully reporting his Bayesian updated belief. However, not in order to hide the truth, but to help it come out at the end: to prevent the process from settling into a commonly known belief—the "Aumann conditions"—on a certain subset of the set of possible states of the world (in which the process then would be blocked), and this way make it converge to a subset of the set of possible states of the world on which it will be commonly known whether the event in question has occurred or not. The strategic movement described in this example is similar to a conversational implicature: the correct interpretation of the deviation from truthfully reporting the Bayesian updated belief thrives on it being common knowledge that the announced probability cannot possibly be the speaker's Bayesian updated belief at this step. Finally, the argument is embedded in a game-theoretic model.

32 – Andrew Charman – *The Census and the Second Law: An Entropic Approach to Optimal Apportionment of Legislative Representatives*

Abstract: Apportionment of representatives and legislative re-districting represent reversals of the usual democratic dynamic, since politicians are in effect choosing their constituents instead of having voters choose their representatives. In the United States, debate over how best to convert census data into congressional seats extends back to the country's founding and persists to the present day. Borrowing ideas of entropy maximization that underly both statistical thermodynamics and information theory, a method is derived for optimal apportionment of representatives from the various states to the U.S. House of Representatives. The very same variational principle, and associated "greedy" minimization algorithm, can also be used to assess or choose the optimal size of the House, and after re-districting following reapportionment, to assess differences in Congressional district sizes. The ideal apportionment of seats to states is that which minimizes the relative entropy (subject to constraints on the total number of representatives, as well as on the allowed minimum and maximum number of representatives per state), or equivalently, maximizes the entropy of individual weights of representation.
The Kullback-Leibler divergence (i.e., relative entropy) is the only natural measure of discrepancy between actual "shares of the vote" held by individuals under a proposed

apportionment, and ideal shares based on the goal of proportional representation.
A probabilistic interpretation can be given to voting shares in order to invoke the characterization theorems leading uniquely to the entropy functional, though other axiomatizations (such as those of Skilling and Knuth) will apply to any nonnegative additive measures, so can justify the approach even if the probabilistic reading is questioned.
While we focus on the case of the U.S. Congress, for which the entropic apportionment disagrees with the currently adopted apportionment method when applied to the 2020 U.S. Census results, the same principle could be applied to other examples or levels of representative government, including allocation of seats in many types of federal, parliamentary, or party-list electoral systems—or more broadly, to any constrained optimization problem where integral or indivisible rewards are to be allocated based on fractional dessert.

12:20 – 13:20 **Lunch break**

13:20 – 15:00 **Session 3: Imaginging Applications**

Chaired by: Ali Mohammad-Djafari / Kevin Knuth

52 - Invited Speaker 03 : Lorenzo Valzania - ***Computational tools to control light through complex media***

Abstract: Materials multiply scattering and diffusing light, also known as complex media, are found in a wide variety of systems, like the atmosphere, biological tissues and multimode fibers. This makes scattering in complex media a prolific research topic. Wavefront shaping techniques have established themselves as the tools of choice to guide light through complex media. A current challenge lies in the stability times of living specimens, which allow light control for up to only a few ms. This sets severe constraints on the required frame rates which, in turn, results in measurements with inherently low signal-to-noise ratios.
In the first part of my talk, the most widespread wavefront shaping algorithms and their range of applicability will be introduced. To lift their current limitations, I will present an optimization-based method which relies on adaptive filter theory. Using a recursive least-squares (RLS) algorithm, I estimate the linear response (also known as transmission matrix, TM) of a scattering medium in an online and recursive manner. Crucially, this makes the technique ideally suited for long streams of data, as the computational complexity is independent of the amount of measurements. The least-squares optimization ensures optimal resilience to noise. Finally, the algorithm is provided with a tunable memory, such that the estimation of the TM adapts to the stability time of the medium and tracks its dynamics. I will showcase its performance on three tasks, namely focusing, point-spread function engineering, and maximal energy transmission through

dynamic scattering media, at realistic noise levels and stability times.
The second part of my talk will revolve around a longstanding yet challenging problem in physics and signal processing, namely phase retrieval. The non-linear equation of phase retrieval arises whenever one deals with reconstructions from phaseless linear projections, and appears in many different scenarios, from computational imaging to optical computing. I will illustrate how the study of phase retrieval benefits from the recent advances in machine learning, and vice versa. On the one hand, the analogy between phase retrieval and single-layer neural networks has improved our understanding on the behavior of neural networks. On the other hand, deep-learning regularization opens up new possibilities for phase retrieval applications and computational imaging. Focusing on computational imaging applications, I will review phase retrieval under a unifying framework and provide an overview of phase retrieval algorithms, from flexible gradient descent routines to more specialized spectral methods. In particular, I will stress how phase retrieval with random projections enjoys strong convergence guarantees, and is naturally implemented when imaging through scattering.

31 - Berlin Chen and Cyrus Mostajeran - *Geometric learning of hidden Markov models via a method of moments algorithm*

Abstract: We present a novel algorithm for learning the parameters of hidden Markov models (HMMs) in a geometric setting where the observations take values in Riemannian manifolds. In particular, we elevate a recent second-order method of moments algorithm that incorporates non-consecutive correlations to a more general setting where observations take place in a Riemannian symmetric space of non-positive curvature and the observation likelihoods are Riemannan Gaussians. The resulting algorithm decouples into a Riemannian Gaussian mixture model estimation algorithm followed by a sequence of convex optimization procedures. We demonstrate through examples that the learner can result in significantly improved speed and numerical accuracy compared to existing learners.

79 - Yannis Kalaidzidis - *Revisiting multiple signal classification algorithm for super-resolution fluorescence microscopy in presence of non-uniform background*

Abstract: Single Molecule Localization Microscopy (SMLM) achieves super-resolution by utilizing photoswitchable fluorescent labels to separate signals from individual molecules in time and space. In this case, the localization of each individual fluorophore (the position of the center of the diffraction spot) can be found with an accuracy of 10÷50 nanometers. Drawback of SMLM is the limitation of the low labelling density of the biological samples and need to acquire thousands of time-lapse frames for one final image. The first constraint ensures the spatial separation of signals from fluorophores at each time frame, while the second allows a sufficient number of localizations to be collected for the final image reconstruction. The multiple signal classification algorithm (MUSICAL) provides a mean to overcome both limitations and reconstruct super-resolution image from a densely labelled sample with few tens of time-lapse frames.

In short, MUSICAL uses singular value decomposition (SVD) to separate spatially overlapping fluorophore signals with independent temporal profiles and the projection of the microscope point spread function (PSF) onto null-space of the SVD spatial matrix as an indicator function for localization. Unfortunately, MUSICAL is limited to images with uniform low-level backgrounds and requires user-defined parameter for separation of kernel- and null-spaces of SVD spatial matrix. In presented work, a probabilistic method for separating kernel- and null-spaces of SVD spatial matrix is proposed, as well as a method for probabilistic estimating the contribution of the inhomogeneous background to the first column of the SVD spatial matrix. The performance of modified MUSICAL is demonstrated on simulated and experimental data.

74 – Geoffroy Delamare and Ulisse Ferrari - *Time-Dependent Maximum Entropy Model for Populations of Retinal Ganglion Cells*

Abstract: The inverse Ising model is used in computational neuroscience to infer probability distributions of the synchronous activity of large neuronal populations. This method allows for finding the Boltzmann distribution with single neuron biases and pairwise interactions that maximizes the entropy and reproduces the empirical statistics of the recorded neuronal activity. Here we apply this strategy to large populations of retinal output neurons (ganglion cells) of different types, stimulated by multiple visual stimuli with their own statistics. The activity of retinal output neurons is driven by both the inputs from upstream neurons, which encode the visual information and reflect stimulus statistics, and the recurrent connections, which induce network effects. We first apply the standard inverse Ising model approach, and show that it accounts well for the system's collective behavior when the input visual stimulus has short-ranged spatial correlations, but fails for long-ranged ones. This happens because stimuli with long-ranged spatial correlations synchronize the activity of neurons over long distances. This effect cannot be accounted for by pairwise interactions, and so by the pairwise Ising model. To solve this issue, we apply a previously proposed framework that includes a temporal dependence in the single neurons biases to model how neurons are driven in time by the stimulus. Thanks to this addition, the stimulus effects are taken into account by the biases, and the pairwise interactions allow for characterizing the network effect in the population activity and for reproducing the structure of the recurrent functional connections in the retinal architecture. In particular, the inferred interactions are strong and positive only for nearby neurons of the same type. Inter-type connections are instead small and slightly negative. Therefore, the retinal architecture splits into weakly interacting subpopulations composed of strongly interacting neurons. Overall, this temporal framework fixes the problems of the standard, static, inverse Ising model and accounts for the system's collective behavior, for stimuli with either short or long-range correlations.

15:00 – 15:20 **Coffee break**

15:20 – 16:40

Session 4: Bayesian Simulation And Model Selection

Chaired by: Geert Verdoolaege / Pierre-Henri Wuillemin

38 - Florent Leclercq - *Simulation-based inference of Bayesian hierarchical models while checking for model misspecification*

Abstract: This paper presents recent methodological advances to perform simulation-based inference (SBI) of a general class of Bayesian hierarchical models (BHMs), while checking for model misspecification. Our approach is based on a two-step framework. First, the latent function that appears as second layer of the BHM is inferred and used to diagnose possible model misspecification. Second, target parameters of the trusted model are inferred via SBI. Simulations used in the first step are recycled for score compression, which is necessary to the second step. As a proof of concept, we apply our framework to a prey-predator model built upon the Lotka-Volterra equations and involving complex observational processes.

70 - Orestis Loukas and Ho Ryun Chung - *Model selection in the world of Maximum Entropy*

Abstract: Science aims at identifying suitable models that best describe a population based on a set of features. Lacking information about the relationships among features there is no justification to a priori fix a certain model. Ideally, we want to incorporate only those relationships into the model which are supported by observed data. To achieve this goal the model that best balances goodness of fit with simplicity should be selected. However, parametric approaches to model selection encounter difficulties pertaining to the precise definition of the invariant content that enters the selection procedure and its interpretation. A naturally invariant formulation of any statistical model consists of the joint distribution of features, which provides all the information that is required to answer questions in classification tasks or identification of feature relationships. The principle of Maximum Entropy (MaxEnt) offers a framework to directly estimate a model for this joint distribution based on phenomenological constraints. Reformulating the inverse problem to obtain a model distribution as an under-constrained linear system of equations, where the remaining degrees of freedom are fixed by entropy maximization, tremendously simplifies large-N expansions around the optimal distribution of Maximum Entropy. We have exploited this conceptual advancement to clarify the nature of prominent model-selection schemes providing an approach to systematically select significant constraints evidenced by the data. To facilitate the treatment of higher-dimensional problems, we propose HyperMaxEnt - a clustering method to efficiently tackle the MaxEnt selection procedure. We demonstrate the utility of our approach by applying the advocated methodology to analyze long-range interactions from spin glasses and uncover three-point effects in COVID-19 data.

67 - Gordon Morison and Adam Brown - ***Bayesian statistics approach to imaging of aperture synthesis data: RESOLVE meets ALMA***

Abstract: The Atacama Large Millimeter/submillimeter Array (ALMA) is currently revolutionizing observational astrophysics. The aperture synthesis technique provides angular resolution otherwise unachievable with the conventional single-aperture telescope.
However, recovering the image from the inherently undersampled data is a challenging task. CLEAN algorithmhas proven successful and reliable and is commonly used in imaging the interferometric observations. It is not, however, free of limitations. Point-source assumption, central to the CLEAN is not optimal for the extended structures of molecular gas recovered by ALMA. Additionally, negative fluxes recovered with CLEAN are not physical. This begs to search for an alternatives that would be better suited for specific science cases.
We present the recent developments in imaging ALMA data using Bayesian inference techniques, namely the RESOLVE algorithm. This algorithm, based on information field theory has been already successfully applied to image the Very Large Array data.
We compare the capability of both CLEAN and RESOLVE to recover known sky signal, convoluted with the simulator of ALMA observation data and we investigate the problem with a set of actual ALMA observations.

17 - Riko Kelter - ***The Full Bayesian Evidence Test: Theory and Applications in Bayesian response-adaptive clinical Trial Designs***

Abstract: The Full Bayesian Evidence Test (FBET) has recently been developed as a generalization of the Full Bayesian Significance Test (FBST) and advocates the use of the Bayesian evidence value to test a hypothesis statistically. The FBET builds on the Bayesian evidence interval, which is a generalization of existing Bayesian interval estimates and is motivated from the principle of least surprise which was first proposed by I.J. Good. Although the existing theory is promising, it offers a wide range of choices how to employ the FBET in practice.
In this tutorial paper, the theory underpinning the Bayesian evidence interval, FBET and Bayesian evidence value is outlined and guidance is provided how to perform a hypothesis test with the FBET in practice. Therefore, a recently developed software implementation in the statistical programming language R is used to provide illustrating examples for testing the hypothesis of drug efficacy in Bayesian response-adaptive clinical trials designs] via the Bayesian evidence value. Next to stopping early for efficacy or futility in adaptive clinical trials, testing an interval hypothesis is illustrated using data from medical research. Possible extensions and future research with a focus on maximum entropy are discussed.

16:40 – 17:00 **Coffee break**

17:00 – 18:20 **Session 5: Quantum Systems**

Chaired by: Udo Von Toussaint / Maria Trocan

19 – Philipp Joppich, Sebastian Dorn, Oliver De-Candido, Wolfgang Utschick and Jakob Knollmüller - *Classification and Uncertainty Quantification of Corrupted Data using Semi-Supervised Autoencoders*

Abstract: Parametric and non-parametric classifiers often have to deal with real-world data, where corruptions like noise, occlusions, and blur are unavoidable. We present a probabilistic approach to classify strongly corrupted data and quantify uncertainty, even though the corrupted data does not have to be included to the training data.
A supervised autoencoder is the underlying architecture. We use the decoding part as a generative model for realistic data and extend it by convolutions, masking, and additive Gaussian noise to describe imperfections. This constitutes a statistical inference task in terms of the optimal latent space activations of the underlying uncorrupted datum. We solve this problem approximately with Metric Gaussian Variational Inference (MGVI). The supervision of the autoencoder's latent space allows us to classify corrupted data directly under uncertainty with the statistically inferred latent space activations. We show that the derived model uncertainty can be used as a statistical «lie detector» of the classification. Independent of that, the generative model can optimally restore the corrupted datum by decoding the inferred latent space activations.

27 – Philippe Jacquet - *Is Quantum Tomography a difficult problem for Machine Learning?*

Abstract: One of the key problem in machine learning is the characterization of the learnability of a problem. The regret is way to quantify learnability. The quantum tomography is a special case of machine learning where the training set is a set of quantum measurements and the ground truth the results of these measurements, but nothing is known about the hidden quantum system. We will show that in some case the quantum tomography is hard to learn problem.
We consider a problem related to optical fiber communication where information are encoded in photon polarizations. The problem is to determine the polarisation angles supposed to be unknown by the receiver and its determination. For this purpose, the sender sends a sequence of T equally polarized photons, the receiver measures these photons over a collection of T measurement angles. The labels are the sequence of binary measurement obtained.
This problem is the most simplified version of tomography on quantum telecommunication, since it relies on a single parameter. More realistic and complicated situation will occur when information are embedded in circular polarization within groups of photons. This will considerably increase the dimension of the feature vectors and make more relevant our results on training process. However in the situation analyzed in our paper, we show that this simplistic system is difficult to learn.

82 - Mariela Portesi, Yanet Alvarez, Marcelo Losada and Gustavo Martin Bosyk - *Coherence-impurity complementarity for quantum systems using entropies and majorization*

Abstract: Singh et al. studied the bound imposed by mixedness on the amount of quantum coherence present in a quantum system. In particular, they found a nontrivial bound for the sum of coherence and mixedness when they are given, respectively, by the L1-norm and the linear entropy. Moreover, they showed a trade-off between the relative entropy of coherence and von Neumann entropy.
Based on those results, coherence-impurity complementarity is expected to be a general characteristic of quantum systems, regardless of the measures used. We study general quantifiers of coherence and impurity, namely: the generalized coherence vector and the vector of eigenvalues, respectively. We observe the existence of a nontrivial bound for the tensor product between both vectors, which captures the trade-off between coherence and mixedness in general. In particular we obtain the optimal bound in the case of a single-qubit system, and approximate numerical bounds for the qutrit and for dimension 4.

34 - François Verdeil, Yannick Deville and Alain Deville - *A unitary quantum process tomography algorithm robust to systematic errors*

Abstract: Quantum process tomography (QPT) methods aim at identifying a given quantum process. QPT is a major quantum information processing tool, since it especially allows one to characterize the actual behavior of quantum gates, which are the building blocks of quantum computers. The present paper focuses on the estimation of a unitary process. This class is of particular interest because quantum mechanics postulates that the evolution of any closed quantum system is described by a unitary transformation. Unitary processes have significantly fewer parameters than general quantum processes ($2^{2n_{qb}}$ vs $2^{4n_{qb}}$ - $22n_{qb}$ real independent parameters for nqb qubits).
By assuming that the process is unitary we develop a method that scales better with the size of the system. In the present paper, we stay as close as possible to the standard setup of QPT: the operator has to prepare copies of different input states. The properties those states have to satisfy in order for our method to achieve QPT are very mild. Therefore we choose to operate with copies of $2^{2n_{qb}}$ possibly unknown pure input states. In order to perform QPT without knowing the input states, we perform measurements on half the copies of each state, and let the other half be transformed by the system before measuring them (each copy is only measured once). This setup has the advantage of removing the issue of systematic (i.e. same on all the copies of a state) errors entirely because it does not require the process input to take predefined values. We develop a closed-form solution that first estimates the states from the averaged measurements and then finds the unitary matrix (representing the process) coherent with those estimates by using a closed-form solution of an extended version of Wahba's problem. This estimate is then used as an initial point for a fine tuning algorithm that maximizes the likelihood of the measurements. Simulation results show the effectiveness of the proposed method.

19:00 – 22:00 **Gala dinner at the Procope restaurant**

Wednesday, July 20th

09.00-09.20	**Invited Speaker 04** Will Handley (69)
09.20-09.40	
09.40-10.00	**46** - Johannes Buchner
10.00-10.20	**84** - Nancy Paul
10.20-10.40	**65** - Aleksandr Petrosyan
10.40-11.00	**Coffee Break**
11.00-11.20	**Invited Speaker 05** Livia Partay (02)
11.20-11.40	
11.40-12.00	**50** - Lune Maillard
12.00-12.20	**36** - Alessandra Del Masto
12.20-12.40	**Lunch Break**
12.40-13.00	
13.00-13.20	
13.20-13.40	**Invited Speaker 06** Bobak Toussi Kiani (08)
13.40-14.00	
14.00-14.20	**20** - Eliot Tron
14.20-14.40	**21** - Pierre-Yves Lagrave

14.40-15.00	**49** - Sho Sonoda
15.00-15.20	**Transfer to SCAI**
15.20-15.40	**Posters Session** SCAI Sorbonne: 12, 18, 26, 33, 45, 54, 56, 59, 61, 64, 66, 72, 77, 78, 89, 90, 94
15.40-16.00	
16.00-16.20	
16.20-16.40	
16.40-17.00	
17.00-17.20	**SCAI Presentation** & Drinks
17.20-17.40	
17.40-18.00	

Legend:

Topic 4 Bayesian and Maximum Entropy in real world applications

Topic 2 Bayesian Geometric-Informed, Physics-Informed & Thermodynamics-Informed Machine Learning

09:00 – 10:40 **Session 6: Nested Sampling & Applications**

Chaired by: John Skilling / Martino Trassinelli

69 – Invited Speaker 04: Will Handley - *Frontiers in Nested Sampling*

Abstract: In this talk I will give an overview of the state of the art and future research directions in the field of nested sampling. Nested sampling is a numerical method for (a) scanning and optimising a priori unknown functions (b) generating samples from probability distributions, and is uniquely capable of (c) computing high-dimensional numerical integrals and partition functions using probabilistic integration. Implementations of John Skilling's nested sampling meta-algorithm include MultiNest, PolyChord, NeuralNest, DNes, UltraNest & dynesty. Whilst most of this headline literature implements variations on the mechanism for sampling from a likelihood-constrained prior, in the past two decades there have also been several advances in the content of the meta-algorithm and surrounding ecosystem of theory and numerical techniques. These include; dynamic nested sampling, diffusive nested sampling, quantifying parameter estimation errors, post-processing cross-checks using nestcheck and anesthetic, computing p-values in frequentist analyses and using normalising flows to specify priors from previous posterior samples. In my overview of future frontiers, I will discuss reversible nested sampling, using insertion indices or likelihood values rather than likelihood ordering for improved accuracy, applying nested sampling in the field of likelihood-free and simulation-based inference, estimating the end of nested sampling, importance nested sampling, transdimensional nested sampling & multi-objective nested sampling. Interest in collaborations for developing further any of the above will be warmly welcomed.

46 – Johannes Buchner - *Comparison of step samplers for nested sampling*

Abstract: Bayesian inference with nested sampling requires a likelihood-restricted prior sampling method, which draws samples from the prior distribution that exceed a likelihood threshold. For high-dimensional problems, Markov Chain Monte Carlo derivatives have been proposed, based on Metropolis random walks, Hamiltonian Monte Carlo derivatives and slice sampling. We numerically study six slice sampling direction proposals in ellipsoidal, non-ellipsoidal and non-convex problems from 2 to 100 dimensions. The mixing capabilities are vetted with the nested sampling shrinkage test, which makes the results valid regardless how heavy-tailed the posteriors are.
Given the same number of steps, slice sampling prevents collapse of the live points into a linearly dependent sub-space more successfully than hit-and-run sampling. Proposing along differential vectors of live point pairs captures local and non-local geometry, leading to the highest efficiencies.The tested proposals are implemented in the UltraNest (https://johannesbuchner.github.io/UltraNest/) nested sampling package, enabling efficient low and high-dimensional inference of a large class of practical inference problems relevant to astronomy, cosmology, particle physics and astronomy.

84 - Nancy Paul - *Precision Highly Charged Ion Spectroscopy and Atomic Form Factor Studies with Bayesian Analysis*

Abstract: Precision measurements of transition energies in highly-charged ions are important for testing quantum electrodynamics, providing X-ray standards for applications, and understanding astrophysical plasmas. We have made highest precision measurements of M1 transitions in He-like sulfur and argon with the "SIMPA" ECRIS source and the Paris Double Crystal Spectrometer (DCS). These represent the most accurate measurements of these transitions allowing us to test bound-state quantum electrodynamics and relativistic effects in electron-electron correlations. Additionally, as these transitions have extremely narrow natural linewidths (≥ 10^(-7) eV), they are very sensitive to the response function of our instrument. Aided by an ab initio simulation of our spectrometer response function and Bayesian model selection, we have revealed a sensitivity to the underlying atomic form factors describing the response of our silicon crystals to the incident radiation. A systematic evaluation of the model probability based on the evaluation of Bayesian evidence for tens of thousands of different combinations of atomic form factors and doppler broadening widths was performed using the program nested_fit, revealing a new sensitivity to understand systematic effects in our experiment never before studied, potentially a new way to study atomic form factors at low energy.

65 - *Aleksandr Petrosyan - SuperNest: accelerated nested sampling applied to astrophysics and cosmology*

Abstract: We present a method for improving the performance of nested sampling as well as its accuracy. Building on previous work by, we show that posterior repartitioning may be used to reduce the amount of time nested sampling spends in compressing from prior to posterior if a suitable ``proposal" distribution is supplied. We showcase this on a cosmological example, and release the code as an LGPL licensed, extensible Python package supernest (https://gitlab.com/a-p-petrosyan/sspr).

10:40 – 11:00 **Coffee break**

11:00 – 12:20 **Session 7: Applications In Solid State Physics**

Chaired by: Will Handley / Martino Trassinelli

2 - Invited Speaker 05: Livia Partay - *Nested sampling for materials*

Abstract: In recent years we have been working on adapting the Bayesian statistical approach, nested sampling, for studying atomistic systems. Nested sampling automatically generates all the relevant atomic configurations, unhindered by high barriers, and one of its most appealing advantages is that the global partition function can be calculated very

easily, thus thermodynamic properties, such as the heat capacity or compressibility becomes accessible. Nested sampling is a top-down sampling technique, in the materials application this means starting from the high energy region of the potential energy landscape (gas phase configurations) and progressing towards the ground state structure (crystalline solid) through a series of nested energy levels, estimating the corresponding phase space volume of each. This way the method samples the different basins proportional to their volume, and instead of providing an exhaustive list of the local minima, it identifies the thermodynamically most relevant states without any prior knowledge of the structures or phase transitions.
This means that unlike other methods, nested sampling may be fully automated, allowing high-throughput calculations of novel materials. The use and advantages of the nested sampling method has been demonstrated in sampling the potential energy landscape of model systems, often resulting in the discovery of previously unknown thermodynamically stable solid phases, and calculating the pressure-temperature phase diagram of several metals and alloys.

50 - **Lune Maillard** - *Nested sampling for the exploration of potential energy surfaces*

Abstract: In material science, one area of interest is the exploration of the potential energy landscape E which is a function of the position of the atoms. However, when the energy E is derived from calculations where the electrons are treated explicitly (for example by using DFT ...) or when quantum effects are included on the nuclear degrees of freedom, the cost of these explorations increases considerably. Our goal is therefore to reduce the number of sampling points. To do so, we use nested sampling, which turns the multi-dimensional problem into a one-dimensional integral, thus considerably reducing the computational needs for sampling. For this purpose, new exploration methods have been implemented in Nested fit in addition to the random walk, initially implemented for the search of new sampling points (live points). The recognition via newly implemented machine learning methods of cluster structures of the live points, corresponding to the function minima, allow to focus on specific regions that mostly contribute to the energy landscape, considerably improving nested sampling algorithm efficiency. Here, we aim to compare the different methods for searching new sampling points on two benchmark cases in a classical setting: the harmonic potential and the Lennard - Jones atomic clusters.

36 - **Alessandra Del Masto** - *Bulk and point defect properties in α-Zr: uncertainty quantification on a semi-empirical potential*

Abstract: Modelling studies of irradiation defects in α-Zr, such as point defects and their multiple clusters, often use semi-empirical potentials because of their higher computer efficiency as compared to ab initio approaches. Such potentials rely on a fixed number of parameters that need to be fitted to a reference dataset (ab initio and/or experimental), and their reliability is closely related to the uncertainty associated with their parameters, coming from both data inconsistency and model approximations. In this work, parametric uncertainties are quantified on a Second Moment Approximation (SMA) potential, focusing on bulk and point defect properties in α-Zr. A surrogate model, based on polynomial chaos expansion, is first built for properties of interest computed from atomistics, and simultaneously allowing to analytically

compute the sensitivity indices of the observed properties to the potential parameters. This additional information is then used to select a limited number of materials properties for the Bayesian inference. The posterior probability distributions of the parameters are estimated through two Markov Chain Monte-Carlo (MCMC) sampling algorithms. The estimated posteriors of the model parameters are finally used to estimate materials properties (not used for the inference): in any case, most of the properties are closer to the reference ab initio and experimental data than those obtained from the original potential.

12:20 – 13:20 **Lunch break**

13:20 – 15:00 **Session 8: Machine Learning**

Chaired by: Frédéric Barbaresco / Fabrizia Guglielmetti

8 - Invited Speaker 06: Bobak Toussi Kiani - *Quantum algorithms for group convolution, cross-correlation, and equivariant transformations*

Abstract: Group convolutions and cross-correlations, which are equivariant to the actions of group elements, are commonly used in mathematics to analyze or take advantage of symmetries inherent in a given problem setting. Here, we provide efficient quantum algorithms for performing linear group convolutions and cross-correlations on data stored as quantum states. Runtimes for our algorithms are poly-logarithmic in the dimension of the group and the desired error of the operation. Motivated by the rich literature on quantum algorithms for solving algebraic problems, our theoretical framework opens a path for quantizing many algorithms in machine learning and numerical methods that employ group operations.

20 - Eliot Tron - *Equivariant Neural Networks and Differential Invariants Theory for Solving Partial Differential Equations*

Abstract: In this paper, we present two innovative ways of using Equivariant Neural Networks (ENN) to solve PDEs while exploiting the associated symmetries. We first show that Group-Convolutionnal Neural Networks can be used to generalize Physics Informed Neural Networks to encode generic symmetries. By leveraging on differential invariant theory, we then propose using ENN to approximate differential invariants of a given symmetry group, hence allowing to build symmetry preserving Finite Difference methods without the need to formally derivate corresponding numerical invariantizations. Finally, we illustrate the interest of both approaches on the 2D Heat Equation and show in particular that a set of fundamental differential invariants of the roto-translation group SE(2) can be efficiently approximated by ENN for arbitrary functions by training on simple bivariate polynomials evaluations, allowing to easily build SE(2) symmetry-preserving discretization schemes.

21 - Pierre-Yves Lagrave - ***Adaptive Importance Sampling for Equivariant Group-Convolution Computation***

Abstract: This paper introduces an Adaptive Importance Sampling scheme for the computation of group-based convolutions, a key step in the implementation of Equivariant Neural Networks. By leveraging on Information Geometry to dene the parameters update rule when inferring the optimal sampling distribution, we show promising results for our approach by working with the 2-dimensional rotation group SO(2) and von Mises distributions.

49 - Sho Sonoda - ***Closed-form Expression of Parameter Distribution in Neural Network on Noncompact Symmetric Space***

Abstract: Geometric deep learning is an emerging research direction that aims to devise neural networks on non-Euclidean spaces. In this study, we focus on the noncompact symmetric space X=G/K, which covers several important spaces such as the hyperbolic space and the manifold of symmetric positive definite (SPD) matrices, or the SPD manifold. On those spaces, several neural networks have been developed such as hyperbolic neural networks (HNNs) and SPDNets. Despite the common belief that neural network parameters are a blackbox, we have shown that the distribution of parameters in a Euclidean neural network trained by regularized empirical risk minimization (RERM) converges to the ridgelet transform in an over-parametrized regime. The ridgelet transform is an analysis operator that maps a function f on X to the distribution of parameters, written γ, of a network. In this study, we define a fully-connected layer on a noncompact symmetric space, and derive a closed-form expression of the associated ridgelet transform in a unified manner from the perspective of harmonic analysis on symmetric space.

15:00 – 15:20 **Transfert to SCAI Sorbonne**

15:20 – 17:00 **Poster Session**

Chaired by: Lune Maillard / Maria Trocan / Martino Trassinelli / Pierre-Henri Wuillemin

18 - Ranftl, S., Rolf-Pissarczyk, M., Wolkerstorfer, G., Pepe, A., Egger, J., von der Linden, W, & Holzapfel, G.A. - ***Modelling of aortic dissection with Beta random fields and uncertainty propagation with a Bayesian variational auto-encoder***

Abstract: Aortic dissection is a dangerous disease that is linked to stochastic, spatially heterogeneous degradation of aorta tissue. Here, this degradation is modelled as a Beta random field, i.e. a random field for which every marginal follows a Beta distribution, through suitable combination of auxiliary Gaussian random fields. Based on this, the mechanical stresses in the tissue, in response to a mechanical load, are computed with a finite element method. Due to the stochasticity of the model in the input parameters, it is necessary to propagate these uncertainties to the computed mechanical stresses in order to make the simulations interpretable. The prohibitive computational effort of this latter step makes the uncertainty propagation with direct sampling infeasible, requiring instead a surrogate model to be learnt from a small number of samples of the finite element simulation. The structure of the input data, i.e. random field realizations, and the output data, i.e. mechanical stress fields, define a problem similar to image-to-image (I2I) regression. This I2I problem structure suggests a Convolutional Neural Network as a surrogate. Here, we chose a Bayesian Variational Auto-Encoder (B-VAE). After training with Stein Variational Gradient Descent, the B-VAE can subsequently predict approximate stress field samples from random field samples much faster, and with non-parametric Variational Inference estimate the PDF of the predicted mechanical stress fields. A rigorous Bayesian formulation is presented, and the biological implications of the results are discussed.

33 - Tahereh Najafi, Rosmina Jaafar, Rabani Remli, Wan Asyraf Wan Zaidi, Kalaivani Chellappan - *A Computational Model to Determine Membrane Ionic Conductance Using Electroencephalography in Epilepsy*

Abstract: Epilepsy is a multiscale disease that originates in the cellular scale affecting people of all walks of life. The alteration happening at the cellular level affects the electroencephalogram (EEG) signals acquired on the head surface. Despite valuable efforts in epilepsy studies towards increasing fundamental knowledge of this disease, obtaining the information in cellular scale using non-invasive technique is almost impossible. It is a long-term challenge for brain studies due to the inaccessibility of intracranial in-vitro studies. The primary aim of this study is to linkage fundamentals at cellular level to EEG signals acquired from the surface of head skull. This includes discovering the rate of ionic conductance in computational membrane model in epilepsy subjects; during interictal, compared to the normal subjects in response to intermittent photic stimulation (IPS) as a routine test in clinical EEG study. The research is divided into three phases; Phase 1: Model Implementation, Phase 2: Model Modification Ionic Conductance, Phase 3: Validation via Machine Learning Techniques.

54 - Hamideh Manoochehri, Seyed Ahmad Motamedi, Ali Mohammad-Djafari, Masrour Makaremi and Alireza Vafaie sadr - *Attention guided multi-scale CNN Network for Cervical Vertebral Maturation Assessment from Lateral Cephalometric Radiography*

Abstract: Accurate determination of skeletal maturation indicators is crucial in the orthodontic process. Chronologic age is not a reliable skeletal maturation indicator thus Physicians use bone age. Generally, Bone age assessment in the classical radiographic manual methods is done in two main ways: the hand-wrist radiograph method (HWM) and cervical vertebra maturation (CVM). Since cephalometric radiography usually is used in the orthodontic processes, by using the second method the radiation dose, the cost, and the

diagnosis time can be reduced. Determination of CVM degree remains challenging due to the limited annotated dataset, the existence of significant irrelevant areas in the image, the huge intra-class variances, and the high degree of inter-class similarities.
As far as we know, a few related literatures have worked on this application, [1-3] use hand-crafted features and classic machine learning technics and compare the result. [4] use transfer learning on ResNet-18, MobileNet-v2, ResNet-50, ResNet-101, Inception-v3, and Inception-ResNet-v2 to determine the optimal pre-trained network architecture. [5] reports on the development of a Deep Learning (DL) Convolutional Neural Network (CNN) method to determine (directly from images) the degree of maturation of CVM classified in six degrees propose a DCNN network.
to address prenominated challenges, we proposed a novel supervised learning method with a multi-scale attention mechanism, and also, we incorporate the general diagnostic patterns of medical doctors to classify lateral x-ray images as six cervical vertebrae maturation (CVM) classes. Figure (1) shows an overview of the proposed method. We used the multi-scale attention mechanism to highlight the important features, surpass the irrelevant part of the image and efficiently model long-range feature dependencies. attention mechanism improves both their performance and interpretability in visual tasks including image classification. In this work, we used spatial and channel attention. Our proposed network consists of three branches. The first branch extracts local features, creates attention maps and related mask, the second branch uses the mask to extract discriminate features for classification. And third branch fuse local and global features. The result shows that the proposed method can represent more discriminative features therefore the accuracy of image classification can be increased and can provide efficient object localization.

56 - Zahra Amini Farsani, Volker J. Schmid - *Maximum Entropy Regularization and TLBO in the Medical Imaging Problems*

Abstract: In this study, we investigate the application of maximum entropy technique (MET) and regularization functional in addition with an optimization method named teaching learning based optimization (TLBO), to solve a linear ill-posed inverse problem which normally arises in dynamic contrast-enhanced MRI (DCE-MRI) investigations. For that, the main idea is to estimate an appropriate model for the arterial input function (AIF) which is the contrast agent concentration in blood plasma.
Here, a new algorithm is described to estimate an acceptable probability distribution function for AIF according to the MET with regularization functionals and MET with TLBO too, when applying Bayesian estimation approach to estimate two different pharmacokinetic parameters. Estimating the pharmacokinetic parameters in DCE-MRI investigations, is an urgent need to obtain the precise information about the AIF. Moreover, by using the proposed techniques when analyzing simulated and real datasets of the tumors and cancers data according to pharmacokinetic factors, it shows that using Bayesian inference -that infer the uncertainties of the computed solutions, and specific knowledge of the noise and errors - combined with the regularization functional of the maximum entropy problem, improved the convergence behavior and led to more consistent morphological and functional statistics and results.

45 - Xavier Brouty, Matthieu Garcin - *Maxwell's demon and information theory in market efficiency: A Brillouin's perspective*

Abstract: By using Brillouin's perspective on the Maxwell's demon, we determine a new way to describe investor behaviors in the financial markets. The efficient market hypothesis (EMH) in its strong form states that all information in the market, public or private, is accounted for in the stock price. By simulations in a agent-based model, we show that an informed investor using alternative data, correlated to the time series of prices of a financial asset, is able to act as a Maxwell's demon on financial markets. He is then able to do statistical arbitrage consistently with the adaptive market hypothesis (AMH). A new statistical test of market efficiency provides some insight about the impact of the demon on the market. This test determines the amount of information contained in the series, using quantities which are widespread in information theory such as Shannon's entropy. As in Brillouin's perspective, we observe a cycle: Negentropy->Information->Negentropy. This cycle proves the implication of the investor depicted as a Maxwell's demon in the market with the knowledge of alternative data.

61 - Azadeh Fallah, Ali Mohammad-Djafari - *Variational Bayesian Approximation (VBA): A comparison between three optimization algorithms*

Abstract: We propose that coding and decoding in the brain are achieved through digital computation using three principles: relative ordinal coding of inputs, random connections between neurons, and belief voting.
Due to randomization and despite the coarseness of the relative codes, we show that these principles are sufficient for coding and decoding sequences with error-free reconstruction. In particular, the number of neurons needed grows linearly with the size of the input repertoire growing exponentially. We illustrate our model by reconstructing sequences with repertoires on the order of a billion items.
From this, we derive the Shannon equations for the capacity limit to learn and transfer information in the neural population, which is then generalized to any type of neural network. Following the maximum entropy principle of efficient coding, we show that random connections serve to decorrelate redundant information in incoming signals, creating more compact codes for neurons and therefore conveying a larger amount of information. Henceforth, despite the unreliability of the relative codes, few neurons become necessary to discriminate the original signal without error.
Finally, we discuss the significance of this digital computation model regarding neurobiological findings in the brain and more generally with artificial intelligence algorithms, with a view toward a neural information theory and the design of new digital neural networks.

78 - Grégoire Sergeant-Perthuis,David Rudrauf, Dimitri Ognibene, Olivier Belli, Yvain Tisserand - Valentin Durand de Gevigney - *Curiosity driven exploration with perspective taking*

Abstract: Active Inference is one implementation of the Bayesian Brain Hypothesis for generating behaviors similar to those expected from adaptive agents. It relies on an internal representation of the environment that an agent wants to explore and exploit. Using this world

model, the agent continually updates beliefs about plausible competing internal hypothesis on the state of this environment. Under common sensory limitations, Active Inference relates to Partially Observable Markov Decision Process (POMDP). Curiosity or epistemic value is one of the quantities that come into play when given a principle of how the agent should act. In this paper we will consider a toy model of an agent exploring an environment based on curiosity but with a different flavor on how actions are implemented so that they reflect perspective tacking on the internal representation space of the agent. Although essential for psychology especially for multi-agent social interactions (empathy), it is most often absent of existing models of consciousness and the advantages of perspective taking for cybernetics is not that clear. In the following article we propose to compare exploration driven by curiosity with and without perspective taking on the internal representation space and show how perspective taking based on projective transformations induces drastically different behaviors from non perspective taking.

90 – Yushi Li, Frederik Andersson – *Are Central Bankers Inflation Nutters? A Bayesian MCMC Estimator of the Long Memory Parameter in a State Space Model*

Abstract: Several central banks have adopted inflation targets. The implementation of these targets is flexible; the central banks aim to meet the target over the long term but allow inflation to deviate from the target in the short-term in order to avoid unnecessary volatility in the real economy. In this paper, we propose modeling the degree of flexibility using an ARFIMA model. Under the assumption that the central bankers control the long-run inflation rates, the fractional integration order captures the flexibility of the inflation targets. A higher integration order is associated with a more flexible target. We propose a Bayesian Monte Carlo Markov Chain (MCMC) estimator to estimate the integration order. Applying this estimator to inflation from six inflation-targeting countries, we find that inflation targets are implemented with a high degree of flexibility.

59 – Viktoria Kainz, Céline Boehm, Sonja Utz, Torsten Enßlin – *Upscaling Reputation Communication Simulations*

Abstract: Social communication is omnipresent and a fundamental basis of our daily lives. Especially, due to the increasing popularity of social media, communication flows are becoming more complex, faster and more influential. It is therefore not surprising that in these highly dynamic communication structures, strategies are also developed to spread certain opinions, to deliberately steer discussions or to inject misinformation. The reputation game is an agent-based simulation that uses information theoretical principles to model the effect of such malicious behavior taking reputation dynamics as an example. So far, only small groups of 3 to 5 agents have been studied whereas now we extend the reputation game to larger groups of up to 50 agents also including one-to-many conversations. In this setup, the resulting group dynamics are examined, with particular emphasis on the emerging network topology and the influence of agents' personal characteristics thereon. In the long term the reputation game should thus help to determine relations between the arising communication network structure, the used communication strategies and the recipients' behavior, allowing to identify potentially harmful communication patterns, e.g. in social media.

64 – Elham Taghizadeh and Ali Mohammad-Djafari – *SEIR modeling, simulation, parameter estimation, and its application for Covid-19 epidemic prediction*

Abstract: Mathematical models are used in comparing, planning, implementing, evaluating, and optimizing numerous detection, prevention, therapy, and management programs of epidemies. Epidemiology modeling can contribute to the design and analysis of epidemiological surveys, recommend crucial data that should be collected, determine trends, make general forecasts, and estimate the uncertainty in forecasts. There exist a number of models for infectious diseases; as for compartmental models, starting from the very classical SIR and SEIR models to more extended and advanced proposals. Many research works are reported. They show that the SIS, SIR and SEIR models can replicate the dynamics of various epidemics very well. Meanwhile, these models have also been used for the COVID-19.
For instance, Tang et al. investigated a general SEIR epidemiological model where quarantine, isolation and treatment are also considered. Moreover, there also are alternative methods for modeling COVID-19.
Our main contribution in this paper is to consider SEIR compartmented model in detail, to give the significance of its parameters, to do its calibration to a given data set and to estimate its parameters either from a complete set of data or an available subset of them as is the real case. The estimated parameters can then be used for prediction. For the parameter estimation, we used a nonlinear least squares (NLS) optimisation method and a Bayesian estimation method. The nonlinear least squares method (NLS) is used to numerically approximate a solution for the inverse problem within the least-squares fitting. We will apply the different algorithms to the least-squares parameter fitting problem. These algorithms need, in general, a good initialisation.There are many Bayesian estimators, for example the posterior mean, but its inference needs MCMC sampling methods. The computational cost of such methods is huge for this kind of application. We stand here, with more classical techniques of Chi-squared, AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) for which standard technics of computations are, in general, included in many optimization tools available, for example, in Python language packages. A Jupyter notebook of all the simulations of this paper will be available after its publication on github. Once the parameters are estimated, we use the models for prediction of the spread of the virus and compute the probable number of infection and death of individuals. We show the performances of the proposed method on simulated data. Finally, remarking that the fixed parameter model could not give satisfactory results on real data, we propose a time dependent parameter model. Then, this model is implemented and used on real data.

66 – Vincent Eberle, Philipp Frank, Julia Stadler, Silvan Streit, Torsten Enßlin – *The efficient representation of spatially variant point spread functions in Bayesian imaging algorithms*

Abstract: Bayesian imaging algorithms are becoming increasingly important in e.g. astronomy, medicine and biology. Given that many of these algorithms compute iterative solutions to high-dimensional inverse problems, the efficiency and accuracy of the instrument response representation are of high importance for the imaging process. For this reason,

point spread functions, which make up a large fraction of the response functions of telescopes and microscopes, are usually assumed to be spatially invariant in a given field of view and can thus be represented by a convolution. For many instruments, this assumption does not hold and degrades the accuracy of the instrument representation. Here, we discuss the application of butterfly transforms, which are linear neural network structures whose size is scaling subquadratically with the number of data points. Butterfly transforms are efficient by design, since they are inspired by the structure of the Cooley-Tukey Fast Fourier transform. In this work, we combine them in several ways into butterfly networks, compare the different architectures with respect to their performance and identify a representation that is suitable for the efficient respresentation of an synthetic spatially variant point spread function up to a 1% error.

89 - Andree de Backer, Abdelkader Souidi, Etienne A. Hodille, Emmanuel Autissier, Cécile Genevois, Farah Haddad, Antonin Della Noce, Christophe Domain, Charlotte S. Becquart and Marie France Barthe - *Multiobjective Optimization Of The Nanocavities Diffusion In Irradiated Metals*

Abstract: Materials are exposed to severe damaging conditions in nuclear energy production devices: the fission reactors and the fusion tokamaks with magnetic confined plasma (as ITER). Neutrons are produced and stopped in facing materials. They create defects of the microstructure, degrade thermal and mechanical properties and can favour the retention of radioactive tritium in the case of the nuclear fusion energy. To guarantee the safety, dedicated experiment campaigns and multiscale modelling projects are performed for decades. Still few data exist on the properties of diffusion of the nanocavities formed by coalescence of atomic vacancies.The main reasons are that a significant part of the defects are not visible with microscopes and that the simulations play with many temperature dependant processes leading to the microstructure recovery.
In this paper, we combined (i) a systematic experimental study using Transmission Electron Microscopy of irradiated tungsten samples annealed at different temperatures up to 1800 K (producing nanocavity size distribution); (ii) our Kinetic Monte Carlo model of evolution of the microstructure fed by a large collection of atomistic data (e.g. mobility and stability of each size nanocavities – representing several tens of thousands of input physical parameters) (iii) a Multiobjective Optimization method to inverse the problem and obtain the diffusion of nanocavities, which is a parameter of our model, from the comparison with the experimental observations.
To simplify the multi-objective function, we proposed a projection into the space of the diffusion parameters. It revealed non-dominated solutions: two “valleys” of minima corresponding to the nanocavities density and size objectives respectively and which delimitate the Pareto optimal solution. We considered that these “valleys” indicate the upper and lower limits of the uncertainties on the diffusion parameters which remaining after consideration of the experimental results and the model.
We found that the diffusion temperature of nanocavities can be split in three domains: the mono vacancy and small vacancy clusters for which it can be calculated with atomistic model, the small nanocavities for which our approach is decisive and the nanocavities larger than 1.5 nm for which our results agree with the classical surface diffusion theory.

94 - Luis Gonzalez Miret Zaragoza, Martino Trassinelli - ***Study of the relevance of the nuclear three-body interaction in 0f7/2-shell nuclei via Bayesian statistical analysis***

Abstract: The nuclear interaction, which binds together protons and neutrons in atomic nuclei, has still important open questions. At present, it is generally accepted that two-body interactions are insufficient to explain all properties of atomic nuclei and three-body interactions have been proposed to overcome this problem. In this work, we compare the performance of a two-body interaction model and a two-plusthree model within the context of the nuclear shell model by assigning a Bayesian probability to each hypothesis. To do so, we analyse the energy spectra of Z = 20 isotopes and N = 28 isotones and calculate the Bayesian evidence of each of the models. This is done through the nested fit program that implements the Nestled Sampling algorithm. For both two- and three-body interaction hypotheses, the experimental uncertainty are partially incompatible with the deviation from the models. To overcome this inconvenient, the use of different likelihood functions is discussed.

12 - Haithem Ben Khalifa, Wissem Cheikhrouhou and Guy Schmerber - ***Screening of the synthesis route on the structural, magnetic and magnetocaloric properties of La0.6Ca0.2Ba0.2MnO3 manganite: A comparison between solid-solid state process and a combination polyol process and Spark Plasma Sintering***

Abstract: La0.6Ca0.2Ba0.2MnO3 ceramics are prepared by an original route, combining soft chemistry and Spark Plasma Sintering, within a few minutes at 700 °C and by the solid-state reaction at high temperatureswith an annealing temperature of 1200 °C. We have studied the leverage of the powder synthesismethod on the structural, morphological, magnetic and magnetocaloric properties of the samples. X-ray diffraction analysis using Rietveld refinement revealed that our materials crystallize in the rhombohedral system with R3-c space group for the sample prepared by the Polyol-Spark Plasma Sintering method and in the orthorhombic structure with Pbnm space group for the sample synthesized by the solid-state reaction. Magnetization measurements versus temperature under magnetic applied field of 0.05 T show a paramagnetic-ferromagnetic phase transition for both samples. The Arrott plots reveal that ours materials undergo a second-order phase transition. The maximum values of the magnetic entropy change (-ΔSmax M) under the magnetic field change of 5 T are 2.4 and 4.7 J/kg K for La0.6Ca0.2Ba0.2MnO3 synthesized by using solid-state reaction and Polyol-Spark Plasma Sintering methods respectively. The highest value of the relative cooling power RCP is found to be 244 J/kg for the Polyol-Spark Plasma Sintering sample under 5 T. These results are interesting enough and suggest that the Polyol-Spark Plasma Sintering synthesis method is a feasible route to prepare high quality perovskite material for magnetic cooling application.

26 - Ning Chu, Ali Djafari, Liang Yu and Li Wang - ***Infrared Temperature reconstruction based on Bayesian Varational Approximation***

Abstract: Infrared imaging is widely used to detect abnormal body temperature thanks to its non-contact and large-scale thermography. However, the measured temperature is always much lower than the ground-truth. Such measuring errors are related to many factors, such as the surface emissivity, measuring distance, ambient temperature and humidity etc. The temperature

reconstruction method proposed in this paper has two important points. The first point is to use the improved infrared radiation model (IIRM). The surface emissivity in the IIRM is obtained through empirical Bayesian learning (EBL). The relationship between atmospheric transmittance and air humidity in the IIRM is established, so that the IIRM is suitable for measuring scenarios with different humidity. The second point is to solve the inverse problem based on Bayesian Varational Approximation (VBA). The previous methods to reduce the error of infrared temperature measurement are mainly to improve the performance of the hardware or use regression methods to fit the performance curve of the infrared sensor. The mapping relationship between the measured temperature and the actual temperature can be determined by the performance curve. The IIRM-VBA method has three advantages compared to the previous research. First, the measurement error can be reduced without improving the performance of the hardware. Second,the IIRM is suitable for the measuring scenarios with different measuring distances, different humidity, and different emissivity. The robustness of IIRM-VBA is stronger than the previous method of fitting the performance curve. Third, the problem of measurement uncertainty is considered and solved by VBA. The IIRM-VBA method has been validated to have the small errors (≤0.1≈C) through various experiments of black body and human face.

72 - Mariela Portesi, Juan Manuel Pujol and Federico Holik - *Reciprocity relations for quantum systems based on Fisher information*

Abstract: La0.6Ca0.2Ba0.2MnO3 ceramics are prepared by an original route, combining soft chemistry and Spark Plasma Sintering, within a few minutes at 700 °C and by the solid-state reaction at high temperatureswith an annealing temperature of 1200 °C. We have studied the leverage of the powder synthesismethod on the structural, morphological, magnetic and magnetocaloric properties of the samples. X-ray diffraction analysis using Rietveld refinement revealed that our materials crystallize in the rhombohedral system with R3-c space group for the sample prepared by the Polyol-Spark Plasma Sintering method and in the orthorhombic structure with Pbnm space group for the sample synthesized by the solid-state reaction. Magnetization measurements versus temperature under magnetic applied field of 0.05 T show a paramagnetic-ferromagnetic phase transition for both samples. The Arrott plots reveal that ours materials undergo a second-order phase transition. The maximum values of the magnetic entropy change ($-\Delta S^{max}_{M}$) under the magnetic field change of 5 T are 2.4 and 4.7 J/kg K for La0.6Ca0.2Ba0.2MnO3 synthesized by using solid-state reaction and Polyol-Spark Plasma Sintering methods respectively. The highest value of the relative cooling power RCP is found to be 244 J/kg for the Polyol-Spark Plasma Sintering sample under 5 T. These results are interesting enough and suggest that the Polyol-Spark Plasma Sintering synthesis method is a feasible route to prepare high quality perovskite material for magnetic cooling application.

77 - Li Wang, Puhan Zhao, Ning Chu, Ali Mohammad-Djafari and Liang Yu - *Bayesian fusion of infrared image and visible image with a hierarchical Gaussian mixture model*

Abstract: With the development of remote sensing technologies, image fusion has played an essential role in thermal source monitoring. The fusion of visible image and infrared image is typically used in anomaly detection for industrial machines. An infrared image

contains thermal radiation information of the scene, which can distinguish the target from the background. However, infrared images are generally noisy and have lower resolution. On the contrary, the visible image can provide texture details with higher spatial resolution and sharpness. Infrared and visible images are typical multi-modal data that present different kinds of information about the surveyed scene. Therefore, the fusion of infrared and visible images can integrate the key information of infrared image and visible image into a compound image, containing rich background details and clear targets, and therefore improving the effect of key area extraction and temperature monitoring in the industrial systems.

Current algorithms for fusing visible and infrared images can be divided into four categories: multi-scale transformation based methods, sparse representation based methods, subspace learing-based methods and saliency-based methods. Despite the success of previous works on visible and infrared image fusion, the performance stay confined because of the lack of prior information for the application demands, and generally the generalization performance is relatively poor.

The Bayesian method, by defining an appropriate prior to regularize the objective function, is widely used in signal processing tasks, and it shows great potential in multi-band image fusion issues. However, to the best of our knowledge, the infrared and visible image fusion still lacks the utilization of Bayesian method. In this paper, we propose a Bayesian fusion method for infrared and visible image. In the proposed method, a forward model is designed to represent the acquisition process of infrared image and visible image from an unknown fused image, with uncertainties representing acquisition and modeling errors. Hence the fusion problem is cast into an ill-posed inverse problem. The main contribution of the paper can be summarized as follows. First, a forward model to represent the acquisition of infrared and visible image is proposed, and a joint maximum a posterior (JMAP) algorithm is proposed to solve the fusion problem. Second, a wavelet transformation based hierarchical prior model is proposed to model the piece-wise constant fused image. Third, we proposed a new evaluation metric for the fused image, adapting to the remote sensing application. The proposed method has been validated to generate a fused images with smooth and highlighted target areas and clear contour information, which can improve the reliability of the target automatic detection and thermal source monitoring.

Thursday, July 21st

09.00-09.20	**Invited Speaker 07** Olivier Rioul (35)
09.20-09.40	
09.40-10.00	**76 -** Robert Niven
10.00-10.20	**14 -** Margret Westerkamp
10.20-10.40	**83 -** Roman Belavkin
10.40-11.00	**Coffee Break**
11.00-11.20	**Invited Speaker 08** Antoine Bourget (28)
11.20-11.40	
11.40-12.00	**55 -** Marco Armenta
12.00-12.20	**5 -** George Jeffreys
12.20-12.40	**Lunch Break**
12.40-13.00	
13.00-13.20	
13.20-13.40	**Invited Speaker 09** Pierre-Henri Wuillemin (88)
13.40-14.00	
14.00-14.20	**13 -** Jean-Claude Zambrini
14.20-14.40	**40 -** Noémie Combe
14.40-15.00	**Coffee Break**
15.00-15.20	**43 -** Jean-Pierre Françoise
15.20-15.40	**44 -** Andrew Beckett
15.40-16.00	**57 -** Frederic Barbaresco
16.00-16.20	**25 -** Michel Nguiffo Boyom
16.20-16.40	**Coffee Break**
16.40-17.00	**29 -** Valerie Girardin
17.00-17.20	**39 -** Stefan Behringer
17.20-17.40	**58 -** Carlos Alcalde
17.40-18.00	**Sightseeing tours**
18.00-18.20	
18.20-18.40	
18.40-19.00	
19.00-21.00	

Legend:

Topic 1 Foundations of probability, inference, information, entropy

Topic 3 Information theory, machine learning tools for inverse problems

Topic 5 Geometric Statistical Mechanics/ Physics, Lie Groups Thermodynamics & Maximum Entropy Densities

09:00 – 10:40 **Session 9: Foundations**

Chaired by: Romke Bontekoe / Martino Trassinelli

35 – Invited Speaker 07: Olivier Rioul - *What is Randomness? The Interplay between Alpha Entropies, Total Variation and Guessing*

Abstract: In many areas of computer science, it is of primary importance to assess the randomness of a certain variable X. Many different criteria can be used to evaluate randomness, possibly after observing some disclosed data. A "sufficiently random" X is often described as "entropic". Indeed, Shannon's entropy is known to provide a resistance criterion against modeling attacks. More generally one may consider the Rényi α-entropy where Shannon's entropy, collision entropy and min-entropy are recovered as particular cases $\alpha = 1$, 2 and $+\infty$, respectively. Guess work or guessing entropy is also of great interest in relation to α-entropy.
On the other hand, many applications rely instead on the "statistical distance", a.k.a. total variation distance to the uniform distribution. This criterion is particularly important because a very small distance ensures that no statistical test can effectively distinguish between the actual distribution and the uniform distribution.
We establish optimal lower and upper bounds between α-entropy, guessing entropy on one hand, and error probability and total variation distance to the uniform on the other. In this context, it turns out that the best known "Pinsker inequality" and recent "reverse Pinsker inequalities" are not necessarily optimal. We recover or improve previous Fano-type and Pinsker-type inequalities used for several applications.

76 – Robert Niven – *Fluid Densities Defined from Probability Density Functions, and New Families of Conservation Laws*

Abstract: The mass density, commonly denoted $\rho(x, t)$ as a function of position x and time t [SI units: kg m–3], is considered an obvious concept in physics. It is, however, fundamentally dependent on the continuum assumption, the ability of the observer to downscale the mass of atoms present within a prescribed volume to the limit of an infinitesimal volume. In multiphase systems such as fluid-filled porous media, the definition becomes critical, and has been addressed by taking the convolution $[\rho](x, t) =$, involving integration of a local density $\rho(x + r, t)$ multiplied by a weighting function $w(r, t)$ over the small local volume $V(r, t)$, where $[\cdot]$ is an expectation. The weighting function can be formally identified as the local probability density function (pdf) $p(r|t)$, enabling the construction of physical variables from probabilities. This insight is extended to the family of five pdfs $p(u, x|t)$, $p(x|t)$, $p(u|t)$, $p(u|x, t)$ and $p(x|u, t)$, applicable to fluid elements of velocity u and position x at time t in a fluid flow system. By convolution over a small geometric volume V or small velocimetric domain U, these can be used to define five fluid densities, respectively $\zeta(u, x, t)$, $\rho(x, t)$, д(u, t), $\eta(u, x, t)$ and $\xi(u, x, t)$ (dropping expectation symbols). By consideration of the velocimetric as well as geometric domains, the new densities provide a description of fluid flow of higher fidelity than that provided by $\rho(x, t)$ alone. By multiplying with an appropriate specific density, they can also be used to define

the density of any physical quantity in the fluid. Using these densities within an extended form of the Reynolds transport theorem, we derive 11 tables of integral and differential conservation laws applicable to different parameter spaces, for the eight common conserved quantities (fluid mass, species mass, linear momentum, angular momentum, energy, charge, entropy and probability). The findings considerably expand the set of known conservation laws for the analysis of physical systems.

14 – Margret Westerkamp, Igor Ovchinnikov, Philipp Frank and Torsten Enßlin – *Analysis of dynamical field inference in a supersymmetric theory*

Abstract: Stochastic differential equations appear in science, technology and economics. Therefore, the inference of dynamical fields is of paramount importance. We show that the information field theory for dynamical systems, called dynamical field inference, can be used to infer the evolution of fields in dynamical systems from finite data. The central mathematical object of our investigation is the partition function associated with the inference problem from which any relevant quantity of interest such as the mean and the uncertainty can be obtained. This partition function invokes a Dirac delta function as well as a field-dependent functional determinant, to assure the dynamics and proper normalization. Therefore, the evaluation of the corresponding path integral and thus the inference is impeded.
To tackle this problem, Fadeev-Popov ghosts and a Lagrange multiplier are introduced to calculate the normalization determinant and to represent the Dirac function by an integral. According to the supersymmetric theory of stochastics, the action associated with the partition function has a supersymmetry for those ghost and signal fields. In other words, there exists an exchange operation between bosonic and fermionic fields which leaves the system invariant. In this context, the spontaneous breaking of supersymmetry leads to chaotic behaviour of the system. This is one of the major takeaways of the supersymmetric theory of stochastics.
To demonstrate the impact of this breakdown on the predictability of a systems evolution, we show for the case of idealized linear dynamics that the dynamical growth rates of the fermionic ghost fields, characterized by the respective Lyapunov exponents, impact the uncertainty of the field inference. Finally, by establishing perturbative solutions to the inference problem associated with an idealized non-linear system, using a Feynman diagrammatic expansion, we expose that the fermionic contributions, implementing the functional determinant, are key to obtain the correct posterior of the system.

83 – Roman Belavkin, Panos Pardalos and Jose Principe – *Value of Information in the Binary Case and Confusion Matrix*

Abstract: The simplest Bayesian system used to illustrate ideas of probability theory is a coin and a boolean utility function. To illustrate ideas of hypothesis testing, estimation or optimal control one needs to use at least two coins and a confusion matrix accounting to utilities of four possible outcomes. Here we use such a system to illustrate the main ideas of the value of information theory.

10:40 – 11:00 **Coffee break**

11:00 – 12:20 **Session 10: Quivers Geometry And Neural Networks**

Chaired by: Frédéric Barbaresco / Phillipe Jacquet

28 - Invited Speaker 08: Antoine Bourget - *The Geometry of Quivers*

Abstract: Quivers are oriented graphs which have profound connections to various areas of mathematics, including representation theory and geometry. Quiver representations correspond to a vast generalization of classical linear algebra problems. The geometry of these representations can be described in the framework of Hamiltonian reduction and Geometric Invariant Theory, giving rise to the concept of quiver variety.

In parallel to these developments, quivers have appeared to encode in a natural way certain supersymmetric quantum field theories, in particular when they arise from string theory on singular backgrounds or in the presence of D-branes. The associated quiver variety then corresponds to a part of the moduli space of vacua of the theory. However physics tells us that there exists another natural geometric object associated to quivers, which can be seen as a magnetic analog of the (electric) quiver variety. When viewed from that angle, magnetic quivers are a new tool, developed in the past decade, that helps mathematicians and physicists alike to understand geometric spaces.

In this talk, I will review these developments from both the mathematical and physical perspective, emphasizing the deep dialogue between the two communities.

55 - Marco Armenta - *A Notion of Entropy for the study of Neural Network Training Dynamics using Quiver Representations*

Abstract: It is known that the processing of data done by a neural network on a sample of data during the forward pass can be arranged into a thin quiver representation. It has been proved that these quiver representations contain all the information about the forward propagation of data, that is, the order and structure of the operations, the (pre-) activations of neurons, the output of the network and even information about the backward propagation like the gradient. However, for deep neural networks, it is computationally infeasible to save these quiver representations on computer memory for several samples of data and so the numerical study of how these quiver representations evolve in the moduli space of double framed thin quiver representations for deep networks is intractable. Nevertheless, it has been shown that the output of the network is completely determined by a relatively small matrix obtained from the induced quiver representation of a data sample. This matrix has dimensions the number of output neurons by the number of input neurons, which makes it easy to handle for numerical experiments.

With this mathematical framework, one is able to restate the learning process as a statistical mechanical phenomenon in which the training becomes a sequence of a cloud of points inside a moduli space of (double framed thin) quiver representations, that one

then maps to the space of matrices in which one can perform numerical experiments. From these matrices, we construct a notion of entropy in terms of singular values. We perform an empirical analysis of the training dynamics of small neural networks and exhibit that the evolution of this entropy measure is correlated with the performance of the network. For example, a successful training always decreases the entropy of the system; a bad performance is characterized by a plateau in the evolution of the entropy and overfitting is characterized by an increase in entropy. In all cases, we are able to predict when a training is going to be successful or not based only on our measure of entropy. This could bring useful applications like stopping the training when the entropy makes a plateau (for example, while doing hyperparameter search) to save computing resources or even use the measure of entropy to make predictions instead of using the output of the network. We present experiments on multi-layered perceptrons and different datasets to corroborate our conjectures.

5 - George Jeffreys, Siu-Cheong Lau - *Quantum Finite Automata and Quiver Algebras*

Abstract: We find an application in quantum finite automata for the ideas and results of [JL21] and [JL22]. We reformulate quantum finite automata with multiple-time measurements using the algebraic notion of near-ring. This gives a unified understanding towards quantum computing and deep learning. When the near-ring comes from a quiver, we have a nice moduli space of computing machines with metric that can be optimized by gradient descent.

12:20 – 13:20 Lunch break

13:20 – 17:40 Session 11: Information Geometry And Physics

Chaired by: Ariel Caticha / Ali Mohammad-Djafari

88 - Invited Speaker 09 : Pierre-Henri Wuillemin - *Learning Continuous High-Dimensional Models using Mutual Information and Copula Bayesian Networks*

Abstract: High dimensional probabilistic models are not easy to handle and model. In the field of discrete probabilities, the framework of graphical models has allowed to make important advances, however it is not really the case for continuous distributions in general. A model is however able to change the situation by mixing copula and graphical model. We propose to use this model to learn non-parametric graphical models from continuous observational data using non-parametric estimators based on the Bernstein copula and which are constructed by exploiting the relationship between the mutual information and the entropy of the copula. Since they are underpinned by exactly the same

semantic of probabilistic independence as discrete Bayesian networks, we will also show how they open the way to new applications of high-dimensional continuous models such as a posteriori distribution approximation, model explanation, Pearl causality approach, etc.

13 - Jean-Claude Zambrini - *Hamilton-Jacobi-Bellman equations in Stochastic Geometric Mechanics*

Abstract: A few stochastic approaches to Geometric Mechanics are already known.The new one summarised here attributes a fundamental role to Hamilton-Jacobi-Bellman (HJB) equations. Those will be associated with geometric versions of probabilistic Lagrangian and Hamiltonian mechanics. As in the deterministic case, (stochastic) variational principles on Riemannian manifolds will provide Euler-Lagrange equations, shown to be equivalent to Hamiltonian ones. A stochastic Noether Theorem will follow from the Lagrangian viewpoint and give probabilistic first integrals of the stochastic dynamics, namely martingales. Although our results are more general, the inspiration will come from what is called «Schrödinger's problem» in Mass Transportation theory, as well as from the Hydrodynamical interpretation of quantum mechanics.
Our method will use, in particular, tools of the «Second order differential geometry» due to L. Schwartz and P.A. Meyer. In our perspective, this geometry can be interpreted as a kind of probabilistic counterpart of the canonical quantisation procedure for geometric structures of classical mechanics.Our general framework, however, should also be relevant, besides Optimal Transport, in Machine Learning and other fields where HJB equations play a key role.

40 - Noémie Combe - *Poisson geometry of the statistical Frobenius manifold*

Abstract: New insights on parametric families of probability distributions, related to exponential families have been recently given. Throughout the notion of Frobenius manifolds, a deep relation between the class of exponential statistical manifolds, equipped with flat connection and Topological Field theory was possible. The notion of Frobenius manifolds corresponds to a geometrization of the Witten–Dijkraaf–Verlinde–Verlinde PDE equation. In Yu. Manin's classification, this class of statistical manifolds corresponds to the fourth Frobenius manifolds. We refer to it as the statistical Frobenius manifold. In this work, we prove that this source of Frobenius manifolds is a Poisson manifold (i.e. symplectic with Poisson structures). Following the works of Dubrovin–Novikov, one can define Frobenius manifolds by using equations of hydrodynamical type. Such equations of hydrodynamical type imply the existence of Poisson structures. These new statement lead to connecting this new approach to the works that arose from the theory developed by Koszul–Souriau–Vinberg. In particular, it allows to highlight the deep connection existing between those two different insights.

14:40 – 15:00 **Coffee break**

43 - Jean-Pierre Françoise - *Dynamical systems over Lie groups associated to statistical transformation models*

Abstract: A statistical transformation model consists of a smooth data manifold, on which a Lie group smoothly acts, together with a family of probability density functions on the data manifold parametrized by elements in the Lie group. For such a statistical transformation model, the Fisher-Rao semi-definite metric and the Amari-Chentsov cubic tensor are defined on the Lie group. If the family of probability density functions is invariant with respect to the Lie group action, the Fisher-Rao semi-definite metric and the Amari-Chentsov tensor are left-invariant and hence we have a left-invariant structure of a statistical manifold.
In the present work, the general framework of statistical transformation models are explained. Then, the left-invariant geodesic flow associated to the Fisher-Rao metric is considered for a specific family of probability density functions on a compact semi-simple Lie group which plays the roles of both the data manifold and the parameter space. The corresponding Euler-Poincaré and the Lie-Poisson equations are explicitly found in view of geometric mechanics. Related dynamical systems over Lie groups are also mentioned. A generalization in relation to the invariance of the family of probability density functions is further studied.

44 - Andrew Beckett - *Homogeneous Symplectic Spaces and Central Extensions*

Abstract: We summarise recent work on the classical result of Kirillov that any simply-connected homogeneous symplectic space of a connected group G is a hamiltonian G≈-space for a one-dimensional central extension G≈ of G, and is thus (by a result of Kostant) a cover of a coadjoint orbit G≈. We emphasise that existing proofs in the literature assume that G is simply-connected and that this assumption can be removed by application of a theorem of Neeb. We also interpret Neeb's theorem as relating the integrability of one-dimensional central extensions of Lie algebras to the integrability of an associated Chevalley--Eilenberg 2-cocycle.

57 - Frederic Barbaresco - *Souriau Maximum Entropy Statistical Distribution for Homogeneous Bounded Domains: Poincaré and Siegel Disks*

Abstract: We will introduce Gaussian distribution on the space of Symmetric Positive Definite (SPD) matrices, through Souriau's covariant Gibbs density by considering this space as the pure imaginary axis of the homogeneous Siegel upper half space where Sp (2n,R)/U(n) acts transitively. Gauss density of SPD matrices is computed through Souriau's moment map and coadjoint orbits. We will illustrate the model first for Poincaré unit disk, then Siegel unit disk and finally upper half space. For this example, we deduce Gauss density for SPD matrices.

25 - Michel Nguiffo Boyom - Relevant foliations and topological persistence in statistical manifolds

Abstract: In these times, the computing power of computers is likely to reduce the time gap between innovations in pure mathematic and their market applications. Thus, it is appropriate to draw the attention of practitioners of statistics to the relevance of Differential topology. From the applied statistics viewpoint, the samples of polls are but the leaves of the coarse foliations while the leaves smooth foliations are the samples of the smooth polls. From the point of view of the infinitesimal calculation the former are approximations of the latter. The object of the talk is to introduce relevant foliations which yield topological persistence in the category of statistical manifolds of which the statistical models of measurable sets form a sub-category.

16:20 – 16:40 **Coffee break**

29 - Valerie Girardin - ***On the Asymptotic Classification of Generalized Entropies***

Abstract: The classical tool for studying random sequences is the entropy rate, defined as the time averaged limit of the incresing marginal entropies of the sequence; see Caver and Thomas 1991.
Time averaged rates for most generalized entropy functionals are shown in Ciuperca, Girardin, Lhote 2011 to be either infinite or zero. Nevertheless, averaging by some pertinent sequence leads to meaningful generalized entropy rates in Girardin, Lhote 2015, as soon as the random sequence satisfies a smoothness property called quasi-power property (QPP). The QPP holds for a large class of random sequences including Markov chains and continued fractions.
Non trivial time averaged rates are shown in Girardin, Regnault 2022 to occur only for logarithmic versions of Sharma-Taneja-Mittal entropies, apart for the well-known Shannon and Rényi cases.
The Shannon entropy rate of a countable Markov chain is the sum of the entropy of the transition probabilities weighted by the probability of occurrence of each state according to the stationary distribution of the chain; see Cover and Thomas 1991. It is also known in dynamic system theory to depend on the eigenvalue function of a perturbation of the transition matrix. Similarly, different closed-form expressions are obtained for the rates of Markov chains in Regnault, Girardin, Lhote 2017 and Girardin, Lhote, Regnault 2019.
All in all, a classification of generalized entropy functionals into a few number of exclusive types is obtained. The associated axiomatic-free conditions mainly arise from the asymptotic behavior of the marginal entropy of the random sequence.

39 - Stefan Behringer - ***The Value of Information in Circular Settings***

Abstract: This paper investigates two seemingly different concepts of the Value of Information, the classical Hartley information that has many uses in economics and decision theory and the

Shannon/Stratonovich approach that is inspired by results in thermodynamics in models that are set on the boundary of a circle.
Among other things we generalize the original circle example in to n possible realizations and investigate the limit case when the discrete distribution of the random variable approaches a density and its implications for the two Value of Information concepts.

58 - Carlos Alcalde - *Information Geometry In Phase Space*

Abstract: The purpose of this work is to apply concepts of operational physics and Lie group representation theory to the study of the information geometry of classical phase space measurements. We model the parameter space of the statistical manifold as a homogeneous space of Lie Group automorphisms. Our program, based on ideas from quantum mechanics in phase space, leads us to consider a statistical model of Hamiltonian dynamics in the Koopman representation as a Hilbert bundle over a symplectic manifold. The duality of states and observables in our theory is equivalent to a formulation of classical physics in Hilbert space. Classical observables form a commutative Von-Neumann algebra. A physical interpretation of the Koopman wavefunctions is that it is the signal itself. In this work we propose the construction of probability distributions from Lie group quantum states in the Koopman Hilbert space. In this regard we follow ideas of the formulation of Jean Marie Souriau. Statistical states of dynamical systems are defined as probability distributions in the space of motions. States in signal theory arise from representations of dynamical groups. Like Souriau's quantum states, signal states are defined though functions of positive type. Quantum states in the deformation theory approach are defined through the so-called star-exponential. We briefly discuss the relation between the two.

18:00 – 19:30 **Guided tours (for registered participants only)**

Friday, July 22nd

09.00-09.20	**Invited Speaker 10** Anna Simoni (01)
09.20-09.40	
09.40-10.00	**03 -** Frank Nielsen
10.00-10.20	**41 -** Fabio Di Nocera
10.20-10.40	**42 -** Shlomo Dubnov
10.40-11.00	**Coffee Break**
11.00-11.20	**Invited Speaker 11** Fabrizia Guglielmetti (11)
11.20-11.40	
11.40-12.00	**24 -** Phillip Frank
12.00-12.20	**62 -** Harry Bevins
12.20-12.40	**Lunch Break**
12.40-13.00	
13.00-13.20	
13.20-13.40	**Invited Speaker 12** Piotr Graczyk (73)
13.40-14.00	
14.00-14.20	**37 -** Wolfgang Stummer
14.20-14.40	**15 -** Sascha Ranftl
14.40-15.00	**Coffee Break**
15.00-15.20	**7 -** Filippo Masi
15.20-15.40	**9 -** Beatriz Moya

15.40-16.00	**63 -** P. Conde-Cespedes
16.00-16.20	48 - Roland Preuss
16.20-16.40	**Coffee Break**
16.40-17.00	**85 -** Olivier Peltre
17.00-17.20	**68 -** Adrian J. Guel Cortez
17.20-17.40	**22 -** Romke Bontekoe
17.40-18.00	**Closing Business Meeting**
18.00-18.20	
18.20-18.40	
18.40-19.00	
19.00-21.00	

Legend:

Topic 3 Information theory, machine learning tools for inverse problems

Topic 1 Foundations of probability, inference, information, entropy

Topic 2 Bayesian Geometric-Informed, Physics-Informed & Thermodynamics-Informed Machine Learning

09:00 – 10:40

Session 12: Inference And Entropy

Chaired by: Olivier Rioul / Ali Mohammad-Djafari

1 - Invited Speakers 10: Siddhartha Chib, Minchul Shin and Anna Simoni - *Bayesian Exponentially Tilted Empirical Likelihood to Endogeneity Testing*

Abstract: In this paper we consider Bayesian inference for Instrumental Variables (IV) regression models. These models are characterized by moment restrictions and can be cast within the Exponentially Tilted Empirical Likelihood (ETEL) framework. Because the ETEL function has a well defined probabilistic interpretation and plays the role of a nonparametric likelihood, a fully Bayesian semiparametric framework can be developed. We develop an approach based on marginal likelihoods and model comparison to test whether some covariates are endogenous. We first establish desirable frequentist asymptotic properties of our procedure in both the situations with and without endogeneity. In addition, we study how the strength of the IV affects the asymptotic behaviour of the posterior distribution. Then, we analyse the asymptotic behaviour of the marginal likelihoods and Bayes factors for various degrees of endogeneity. Finally, we consider the case where there is a large number of covariates and/or instruments and introduce a regularising prior distribution.
Our approach is illustrated through numerical studies.

3 - Frank Nielsen - *Information measures and information geometry of the Zeta distributions and related truncated distributions*

Abstract: We consider the zeta distributions which are discrete power law distributions that can be interpreted as the counterparts of the continuous Pareto distributions with unit scale. The family of zeta distributions forms a discrete exponential family with normalizing constants expressed using the Riemann zeta function. We report several information-theoretic measures between zeta distributions and study their underlying information geometry.

41 - Fabio Di Nocera - *Unfolding of relative entropies and monotone metrics*

Abstract: In the proposed work, we discuss the geometric aspects of an unfolding procedure of the space of quantum states and show the form of objects relevant in the field of Quantum Information Geometry in this unfolding space. This procedure is then exploited to recognize structures of Classical Information Geometry underlying the quantum world.

42 - Shlomo Dubnov and Gerard Assayag - *Switching Machine Improvisation Models by Latent Transfer Entropy Criteria*

Abstract: In the paper we will present the theory and some experimental results of switching pre-trained models according to second musical improvisation input. Some examples of machine improvisation systems and their use will be demonstrated.

10:40 – 11:00

Coffee break

11:00 – 12:20

Session 13: Bayesian Methods For Imaging

Chaired by: Maria Trocan / Yannis Kalaidzidis

11 - Invited Speaker 11: Fabrizia Guglielmetti - *Bayesian and Machine Learning Methods in the Big Data era for astronomical imaging*

Abstract: The Atacama Large Millimeter/submillimeter Array (ALMA) is an aperture synthesis telescope consisting of 66 high-precision antennas covering a wavelength range from 3.6 to 0.32 mm (84-950 GHz) and during Full Operations will cover a range from 8.5 to 0.32 mm (35-950 GHz). The ALMA antennas can be positioned in a number of different configurations with longest baselines ranging 0.16-16.2 km, which are crucial in determining the image quality and spatial resolution: at the highest frequencies in the most extended configurations, the spatial angular resolution reaches 5 mas at 950 GHz. Sparse sampling, sky and instrumental responses, pervasive presence of noise increase complexities to the demanding task of image reconstruction.
Currently ALMA is generating 1 TB of scientific data daily. Within the next decade, at least one order of magnitude of increased daily data rate is foreseen. Electronic upgrades (receivers and correlator) are planned to improve ALMA sensitivity and observing efficiency. In terms of imaging products, ALMA will produce single field and mosaic cubes of at least two orders of magnitude larger than the current cube size in the GB regime. Since the number of observed spectral lines at once will be duplicated, advanced algorithms are needed to provide shorter processing time while handling larger images. Additionally, the imaging algorithms must provide robust and reliable results to reduce human intervention.
Using real and simulated data sets, we investigate how to employ Bayesian and Machine Learning techniques to tackle the mentioned challenges. The current results of the performance of the different techniques are discussed in view to the ALMA pipeline developments into the Big Data era.

24 - Phillip Frank - *Geometric Variational Inference and its application to Bayesian imaging*

Abstract: Modern day Bayesian imaging problems in astrophysics as well as other scientific areas often result in non-Gaussian and very high dimensional posterior probability distributions as their formal solution. Efficiently accessing the information contained in such distributions remains a core challenge in modern statistics as on the one hand point estimates such as Maximum a Posteriori (MAP) estimates are insufficient due to the non-linear structure of these problems, while on the other hand posterior sampling methods such as Markov-Chain Monte-Carlo (MCMC) techniques may become computationally prohibitively expensive in such high dimensional settings. To nevertheless enable (approximate) inference in these cases, geometric Variational Inference (geoVI) has recently been introduced as an accurate Variational Inference (VI) technique for non-linear, uni-modal probability distributions. It utilizes the Fisher-Rao information metric (FIM) related to the posterior probability distribution and the Riemannian manifold associated with the FIM to construct a set of normal coordinates in which the posterior metric is

approximately the Euclidean metric. Transforming the posterior distribution into these coordinates results in a distribution that takes a particularly simple form which ultimately allows for an accurate approximation with a normal distribution. A computationally efficient approximation of the associated coordinate transformation has been provided by geoVI, which now enables its application to real world astrophysical imaging problems in millions of dimensions.

62 - Harry Bevins - *Marginal Bayesian Statistics Using Masked Autoregressive Flows and Kernel Density Estimators with Examples in Cosmology*

Abstract: Cosmological experiments often employ Bayesian workflows to derive constraints on cosmological and astrophysical parameters from their data. It has been shown that these constraints can be combined across different probes such as Planck and the Dark Energy Survey and that this can be a valuable exercise to improve our understanding of the universe and quantify tension between multiple experiments. However, these experiments are typically plagued by differing systematics, instrumental effects and contaminating signals, which we collectively refer to as `nuisance' components, that have to be modelled alongside target signals of interest. This leads to high dimensional parameter spaces, especially when combining data sets, with >20 dimensions of which only ~5 correspond to key physical quantities. We present a means by which to combine constraints from different data sets in a computationally efficient manner by generating rapid, reusable and reliable marginal probability density estimators, giving us access to nuisance-free likelihoods. This is possible through the unique combination of nested sampling, which gives us access to Bayesian evidences, and the marginal Bayesian statistics code margarine. Our method is lossless in the signal parameters, resulting in the same posterior distributions as would be found from a full nested sampling run over all nuisance parameters, and typically quicker than evaluating full likelihoods. We demonstrate our approach by applying it to the combination of posteriors from the Dark Energy Survey and Planck.

12:20 – 13:20 **Lunch break**

13:20 – 14:40 **Session 14: Graphical And Geometrical Learning**

Chaired by: Zahra Amini Farsani / Ali Mohammad-Djafari

73 - Invited Speaker 09 : Piotr Graczyk, Hideyuki Ishi and Bartosz Kolodziejek - *Graphical Gaussian models associated to a homogeneous graph with permutation symmetries*

Abstract: We impose the invariance on a graphical statistical model under the natural action of a permutation group preserving the conditional independence structure of the

model. We carry out Bayesian model selection. In this abstract, in order to demonstrate our general results for homogeneous cones, we work on the data set of the examination marks of 88 students in 5 different mathematical subjects reported in Mardia et al. (1979), following Høysjaard and Lauritzen (2008).

37 – Michel Broniatowski and Wolfgang Stummer - *A precise bare simulation approach to generalized entropy maximization*

Abstract: In the fields of information theory, statistics, physics, machine learning, artificial intelligence, signal processing and pattern recognition, many flexibilizations of the Shannon entropy (such as e.g. the Renyi and Tsallis entropies) have become frequently used tools. To tackle corresponding constrained maximization problems by a newly developed dimension-free bare (pure) simulation method, is the main goal of this talk. Almost no assumptions (like convexity) on the set of constraints are needed, within our discrete setup of arbitrary dimension, and our method is precise (i.e., converges in the limit). To illustrate the core of our approach, we present numerous examples.This talk is mainly based on our recent preprint arXiv:2107.01693v1.

15 – Sascha Ranftl - *A connection between probability, physics and neural networks*

Abstract: I illustrate an approach that can be exploited for constructing neural networks which a priori obey physical laws. We start with a simple single-layer neural network but refrain from choosing the activation functions yet. Under certain conditions and in the infinite-width limit, we may apply the central limit theorem, upon which the network output becomes Gaussian. We may then investigate and manipulate the limit network by falling back on Gaussian process theory. It is observed that linear operators acting upon a Gaussian process again yield a Gaussian process. This also holds true for differential operators defining differential equations and describing physical laws. If we demand the Gaussian process, or equivalently the limit network, to obey the physical law, then this yields an equation for the covariance function or kernel of the Gaussian process, the solution of which equivalently constrains the model to obey the physical law. Coming back to the central limit theorem, this in turn suggests that neural networks can be constructed to obey a physical law by choosing the activation functions such that they match this particular kernel in the infinite-width limit. I.e., the activation functions constructed in this way guarantee the neural network to a priori obey the physics, up to the approximation error of non-infinite network width in practise. Simple examples of the harmonic oscillator and/or the homogeneous 1D-Helmholtz equation are discussed and compared to naive kernels and activations.

14:40 – 15:00 **Coffee break**

15:00 – 17:40

Session 15: Learning, Surrogate Models And Information

Chaired by: Robert Niven / Ali Mohammad-Djafari

7 - Filippo Masi and Ioannis Stefanou - *Data- and thermodynamics-driven discovery of state variables and evolution equations*

Abstract: The mechanical behavior of complex materials (e.g., meta-, geo, and bio-materials) is hard to grasp with heuristic constitutive models. To this end, data-driven approaches, mostly based on deep learning, have been rising as promising alternatives. Among them, Thermodynamics-based Artificial Neural Networks (TANN) demonstrated that it is possible to predict the multiscale behavior of complex materials from the universal laws of thermodynamics and advanced deep learning approaches. TANN rely on the theory of internal (state) variables, efficiently and accurately accounting for the thermodynamics of (non-)dissipative processes.
Yet, internal variables are hardly identifiable, in a general way, for the large range of complex materials, usually encountered in engineering problems. For this purpose, we develop new thermodynamics-based dimensionality reduction techniques allowing to discover, for the first time, admissible sets of state variables from the knowledge of those quantities characterizing the microscopic material state. We define the latter as internal coordinates (e.g. displacements, velocities, internal forces).
Evolution laws of the state variables are also needed to properly describe the material behavior. For this purpose, we develop a methodology to discover the evolution equations from the enforcement of the first and second laws of thermodynamics.
The capabilities of the proposed method are demonstrated for the constitutive modeling of complex, multiscale materials, displaying inelastic behavior.

9 - Beatriz Moya, Quercus Hernandez, Alberto Badias, Francisco Chinesta and Elias Cueto - *Thermodynamics of learning physical phenomena*

Abstract: In this work we will present how to efficiently incorporate already existing scientific knowledge to the process of learning physics from data. In particular, we investigate how to incorporate the fulfillment of the laws of thermodynamics.
On one hand, a growing interest has been noticed in our community on the imposition of energy conservation. More than twenty works in the last two years have addressed this imposition, either by employing Lagrangian or Hamiltonian formalisms. However, when we deal with dissipative phenomena, the fulfillment of the second law is far from evident. In our recent works, we cast the problem in the so-called dynamical systems equivalence, by which the learning procedure is seen as equivalent to learning from data an ODE governing the evolution of the phase variables of the problem. Under this rationale, the imposition of a GENERIC structure to the evolution equation ensures the fulfillment of the aforementioned principles of thermodynamics. This imposition of a metriplectic structure is done in our case in a soft way. GENERIC is thus seen as an inductive bias for the learning procedure. Other bias, such as an equivariant structure

in a graph-neural network can also help to achieve better accuracy in the predictions. Combined with reinforcement learning strategies, for instance, our approach is able to learn from video sequences without the need of an explicit training on the phenomena at hand.

63 – Patricia Conde-Cespedes, Guillaume Lachaud and Maria Trocan - *The Role of Entropy in Machine Learning and Applications*

Abstract: Although the term entropy was born in 1864 in the field of thermodynamics to measure the degree of disorder in a system, nowadays it is widely used in machine learning approaches. Indeed, later in 1948, the mathematician Claude Shannon in a seminal paper gave birth to Information theory and introduced the concept of information entropy as a measure of information and uncertainty. Since then, the notion of information entropy, also referred to as \textit{Shannon entropy}, has been crucial to the study of the quantification, storage, and communication of information arising in various fields such as statistical Physics, computer Science, economics etc. Roughly speaking, we can define entropy as the average level of information or uncertainty inherent to a random variable.
In this survey paper, we will present a summary of the use of entropy in different approaches and fields of Machine learning, such as decision trees, the loss function to be optimized in neural networks for classification problems, the use of mutual information in feature selection, the entropy of an image for patch-based classification, the entropy of graph, the use of entropy in text mining, the use of entropy in telecommunication systems, etc.

48 – Roland Preuss and Udo von Toussaint - *Surrogate Modelling Of Ion-Solid Interaction Simulations*

Abstract: Data for complex plasma-wall interactions require long running and expensive computer simulations of codes like EIRENE or SOLPS. At the same time the number of input parameters is large which results in a low coverage of the parameter space. Therefore the efficient navigation in the parameter space benefits from an active learning approach which balances exploration and exploitation in the selection of the next measurement to be taken. Using a Gaussian-process (GP) as emulator a Bayesian adaptive exploration method is presented which jointly optimizes the uncertainty of the surrogate target and an user-selected utility function in a region of interest. In addition, the results are compared to Student-t processes as an alternative surrogate model with improved robustness against outliers.

16:20 – 16:40 **Coffee break**

85 - **Olivier Peltre** - *Local Max-Entropy and Free Energy Principles solved by Belief Propagation*

Abstract: By combining information science and differential geometry, information geometry provides a geometric method to measure the differences in the time evolution of the statistical states in a stochastic process. Specifically, the so-called information length (the time integral of the information rate) describes the total amount of statistical changes that a time-varying probability distribution takes through time. In this work, we outline how the application of information geometry may permit us to create energetically efficient and organised behaviour artificially. Specifically, we demonstrate how nonlinear stochastic systems can be analysed by utilising the Laplace assumption to speed up the numerical computation of the information rate of stochastic dynamics. Then, we explore a modern control engineering protocol to obtain the minimum statistical variability while analysing its effects on the closed-loop system's stochastic thermodynamics.

68 - **Adrian Josue Guel Cortez and Eun-jin Kim** - *Information geometry under the Laplace assumption*

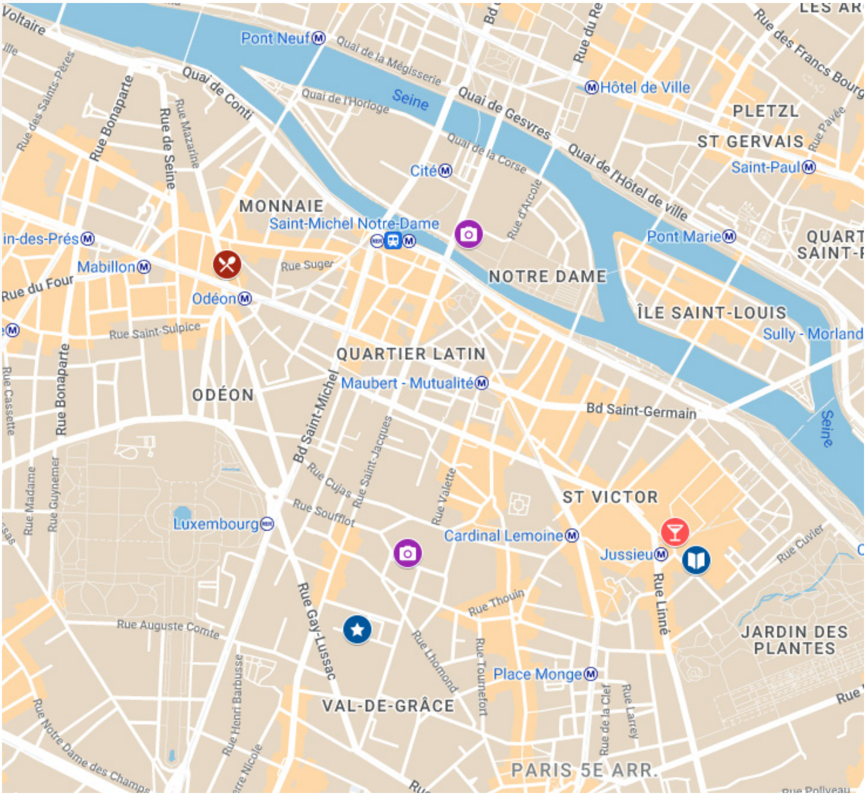
Abstract: By combining information science and differential geometry, information geometry provides a geometric method to measure the differences in the time evolution of the statistical states in a stochastic process. Specifically, the so-called information length (the time integral of the information rate) describes the total amount of statistical changes that a time-varying probability distribution takes through time. In this work, we outline how the application of information geometry may permit us to create energetically efficient and organised behaviour artificially. Specifically, we demonstrate how nonlinear stochastic systems can be analysed by utilising the Laplace assumption to speed up the numerical computation of the information rate of stochastic dynamics. Then, we explore a modern control engineering protocol to obtain the minimum statistical variability while analysing its effects on the closed-loop system's stochastic thermodynamics.

22 - **Romke Bontekoe and Barrie J. Stokes** - *Kangaroos in Cambridge*

Abstract: In this tutorial paper, the kangaroo problem is revisited by two different methods. We compare the classical maximum entropy approach with the information geometry approach.

16:40 – 19:00 **Closing session and business meeting**

MAXENT 2022 EVENT MAP



Conference Location

July 18th to 22nd
IHP - Sorbonne Université
11 Rue Pierre et Marie Curie, 75005 Paris

Poster Session

July 20th
SCAI Sorbonne Campus Pierre et Marie Curie
Place Jussieu, 75005 Paris

Welcome Cocktail

July 18th
SCAI Sorbonne Campus Pierre et Marie Curie
Place Jussieu, 75005 Paris

Gala Dinner

July 19th
Le Procope Restaurant
13 Rue de l'Ancienne Comédie, 75006 Paris

Tour 1: Ile de la Cité

Meeting point:
Notre-Dame Cathedral Square, 75004 Paris

Tour 2: Rue de la Montagne Sainte-Geneviève

Meeting point:
Place du Panthéon, 75005 Paris

Monday July 18	Tuesday July 19	Wednesday July 20	Thursday July 21	Friday July 22
Registration 08:30 - 09:00	**Registration** 08:30 - 09:00			
Tutorial 01 09:00 - 10:00	**Session 1** 09:00 - 10:40	**Session 6** 09:00 - 10:40	**Session 9** 09:00 - 10:40	**Session 12** 09:00 - 10:40
Tutorial 02 10:00 - 11:00	**Coffee break** 10:40 - 11:00	**Coffee break** 10:40 - 11:00	**Coffee break** 10:40 - 11:00	**Coffee break** 10:40 - 11:00
Coffee break 11:00 - 11:20	**Session 2** 11:00 - 12:20	**Session 7** 11:00 - 12:20	**Session 10** 11:00 - 12:20	**Session 13** 11:00 - 12:20
Tutorial 03 11:20-12:20				
Lunch break 12:20 - 13:20	**Lunch break** 12:20 - 13:20	**Lunch break** 12:20 - 13:20	**Lunch break** 12:20 - 13:20	**Lunch break** 12:20 - 13:20
Tutorial 04 13:20 - 14:20	**Session 3** 13:20 - 15:00	**Session 8** 13:20 - 15:00	**Session 11** 13:20 - 14:40	**Session 14** 13:20 - 14:40
Tutorial 05 14:20 - 15:20	**Coffee break** 15:00 - 15:20	**Transfer to SCAI** 15:00 - 15:20	**Coffee break** 14:40 - 15:00	**Coffee break** 14:40 - 15:00
Coffee break 15:20 - 15:40	**Session 4** 15:20 - 16:40	**Poster Session** 15:20 - 18:00	**Session 11** 15:00 - 16:20	**Session 15** 15:00 - 16:20
Tutorial 06 15:40 - 16:40			**Coffee break** 16:40 - 17:00	**Coffee break** 16:40 - 17:00
Coffee break 16:40 - 17:00	**Coffee break** 16:40 - 17:00		**Session 11** 16:40 - 17:40	**Session 15** 16:40 - 17:40
History of probability 17:00 - 18:00	**Session 5** 17:00 - 18:20		**Sightseeing tours** 18:00 - 19:30	**Closing Session** 17:40 - 19:00
Welcome Cocktail 18:00 – 19:30	**Gala Dinner** 19:00 - 22:00			

PROCEEDINGS

To download the MaxEnt 2022 proceedings, flash the QR code below:

Password:

MAXENT2022

NOTES

International Conference on Bayesian and
Maximum Entropy methods in Science and Engineering

41st MaxEnt2022 Conference

JULY 18-22, IHP, PARIS

CONTACT
maxent22@see.asso.fr

ORGANIZER
www.see.asso.fr