Title
Joint Structures and Common Foundation of Statistical Physics, Information Geometry and Inference for Learning

Subject
The conference will deal with the following topics:

- **Geometric Structures of Statistical Physics and Information**
  - Statistical Mechanics and Geometric Mechanics
  - Thermodynamics, Symplectic and Contact Geometries
  - Lie groups Thermodynamics
  - Relativistic and continuous media Thermodynamics
  - Symplectic Integrators

- **Physical structures of inference and learning**
  - Stochastic gradient of Langevin's dynamics
  - Information geometry, Fisher metric and natural gradient
  - Monte-Carlo Hamiltonian methods
  - Varational inference and Hamiltonian controls
  - Boltzmann machine

Dates
26th July to 31st July 2020
Organizers
Frédéric Barbaresco, THALES, KTD PCC, Palaiseau
Silvère Bonnabel, Mines ParisTech, CAOR, Paris
Géry de Saxcé, Université de Lille, LaMcube, Lille
François Gay-Balmaz, Ecole Normale Supérieure Ulm, CNRS & LMD, Paris
Bernhard Maschke, Université Claude Bernard, LAGEPP, Lyon
Eric Moulines, Ecole Polytechnique, CMAP, Palaiseau
Frank Nielsen, Ecole Polytechnique, LIX, Palaiseau & SONY CSL, Tokyo

Scientific Rational
In the middle of the last century, Léon Brillouin in "The Science and The Theory of Information" or André Blanc-Lapierre in "Statistical Mechanics" forged the first links between the Theory of Information and Statistical Physics as precursors.

In the context of Artificial Intelligence, machine learning algorithms use more and more methodological tools coming from the Physics or the Statistical Mechanics. The laws and principles that underpin this Physics can shed new light on the conceptual basis of Artificial Intelligence. Thus, the principles of Maximum Entropy, Minimum of Free Energy, Gibbs-Duhem's Thermodynamic Potentials and the generalization of François Massieu's notions of characteristic functions enrich the variational formalism of machine learning. Conversely, the pitfalls encountered by Artificial Intelligence to extend its application domains, question the foundations of Statistical Physics, such as the construction of stochastic gradient in large dimension, the generalization of the notions of Gibbs densities in spaces of more elaborate representation like data on homogeneous differential or symplectic manifolds, Lie groups, graphs, tensors, ....

Sophisticated statistical models were introduced very early to deal with unsupervised learning tasks related to Ising-Potts models (the Ising-Potts model defines the interaction of spins arranged on a graph) of Statistical Physics. and more generally the Markov fields. The Ising models are associated with the theory of Mean Fields (study of systems with complex interactions through simplified models in which the action of the complete network on an actor is summarized by a single mean interaction in the sense of the mean field).

The porosity between the two disciplines has been established since the birth of Artificial Intelligence with the use of Boltzmann machines and the problem of robust methods for calculating partition function. More recently, gradient algorithms for neural network learning use large-scale robust extensions of the natural gradient of Fisher-based Information Geometry (to ensure reparameterization invariance), and stochastic gradient based on the Langevin equation (to ensure regularization), or their coupling called "Natural Langevin Dynamics”.

Concomitantly, during the last fifty years, Statistical Physics has been the object of new geometrical formalizations (contact or symplectic geometry, ...) to try to give a new covariant formalization to the thermodynamics of dynamic systems. We can
mention the extension of the symplectic models of Geometric Mechanics to Statistical Mechanics, or other developments such as Random Mechanics, Geometric Mechanics in its Stochastic version, Lie Groups Thermodynamic, and geometric modeling of phase transition phenomena.

Finally, we refer to Computational Statistical Physics, which uses efficient numerical methods for large-scale sampling and multimodal probability measurements (sampling of Boltzmann-Gibbs measurements and calculations of free energy, metastable dynamics and rare events, ...) and the study of geometric integrators (Hamiltonian dynamics, symplectic integrators, ...) with good properties of covariances and stability (use of symmetries, preservation of invariants, ...). Machine learning inference processes are just beginning to adapt these new integration schemes and their remarkable stability properties to increasingly abstract data representation spaces.

Artificial Intelligence currently uses only a very limited portion of the conceptual and methodological tools of Statistical Physics. The purpose of this conference is to encourage constructive dialogue around a common foundation, to allow the establishment of new principles and laws governing the two disciplines in a unified approach. However, it is also about exploring new « chemins de traverse ». 
<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>09.00_09.30</td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>09.30_10.00</td>
<td>Talk 1</td>
<td>Manoela de Leon</td>
</tr>
<tr>
<td>10.00_10.30</td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>10.30_11.00</td>
<td>Talk 2</td>
<td>Gery de Saxcé &amp; Charles-Marie Marle</td>
</tr>
<tr>
<td>11.00_11.30</td>
<td>Talk 3</td>
<td>Kevin Grosvenor</td>
</tr>
<tr>
<td>11.30_12.00</td>
<td>Lunch Break</td>
<td></td>
</tr>
<tr>
<td>12.00_12.30</td>
<td>Talk 4</td>
<td>Frédéric Barbaresco &amp; Frank Nielsen</td>
</tr>
<tr>
<td>12.30_13.00</td>
<td>Lunch Break</td>
<td></td>
</tr>
<tr>
<td>13.00_13.30</td>
<td>Talk 5</td>
<td>Géry de Saxcé</td>
</tr>
<tr>
<td>13.30_14.00</td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>14.00_14.30</td>
<td>Talk 6</td>
<td>Eric Moulines</td>
</tr>
<tr>
<td>14.30_15.00</td>
<td>Lunch Break</td>
<td></td>
</tr>
<tr>
<td>15.00_15.30</td>
<td>Talk 7</td>
<td>Arjan van der Schaft</td>
</tr>
<tr>
<td>15.30_16.00</td>
<td>Lunch Break</td>
<td></td>
</tr>
<tr>
<td>16.00_16.30</td>
<td>Talk 8</td>
<td>François Gay-Balmaz</td>
</tr>
<tr>
<td>16.30_17.00</td>
<td>Talk 9</td>
<td>Frédéric Barbaresco</td>
</tr>
<tr>
<td>17.00_17.30</td>
<td>Lunch Break</td>
<td></td>
</tr>
<tr>
<td>17.30_18.00</td>
<td>Talk 10</td>
<td>Géry de Saxcé</td>
</tr>
<tr>
<td>18.00_18.30</td>
<td>Dinner</td>
<td></td>
</tr>
<tr>
<td>18.30_19.00</td>
<td>Talk 11</td>
<td>Giovanni Pistone</td>
</tr>
<tr>
<td>19.00_19.30</td>
<td>Talk 12</td>
<td>Koichi Tojo</td>
</tr>
<tr>
<td>19.30_20.00</td>
<td>Posters</td>
<td></td>
</tr>
<tr>
<td>20.00_20.30</td>
<td>Talk 13</td>
<td>Steve Huntsman</td>
</tr>
<tr>
<td>20.30_21.00</td>
<td>Talk 14</td>
<td>Pratik Chaudhari</td>
</tr>
<tr>
<td>21.00_21.30</td>
<td>Talk 15</td>
<td>Zdravko Terze</td>
</tr>
</tbody>
</table>

**Contact Geometry and Thermodynamical Systems**

**Dinner**

**SSD Jean-Marie Souriau’s Book 50th Birthday**

**GSI’21 Paris 2021 Announcement**

**Information Geometry and Quantum Fields**

**Free Time**

**Diffeological Fisher Metric**

**Schroedinger’s Problem, HJB Equations**

**Langevin Dynamics: Old and News**

**Learning Physics from Data**

**Info. Geometry and Integrable Hamiltonian**

**On Statistical Distances and Information Geometry for Machine Learning**

**Computational Information Geometry**

**Non-Equilibrium Thermodynamic Geometry: A Variational Perspective of Closed & Open Systems**

**Geometric Mechanics: Gallilean Mechanics & Thermodynamics of Continua**

**Geometric Mechanics: Souriau-Casimir Lie Groups**

**Thermodynamics & Machine Learning**

**Thermodynamic Efficiency Implies Predictive Inference**

**Exponential Family by Representation Theory**

**Computational Dynamics of Multibody-Fluid Systems**

**Covariant Momentum Map Thermodynamics**

**Mechanics of the Probability Simplex**

**Sampling and Statistical Physics via Symmetry**

**Deep Learning as Optimal Control**

**The Bracket Geometry of Flows and Diffusions**

**Dirac Structures in Thermodynamics**

**Port Thermodynamic Systems Control**

**Learning with Few Labeled Data**

**Computational Dynamics of Multibody-Fluid Systems**

**Dirac Structures in Thermodynamics**

**Port Thermodynamic Systems Control**

**Learning with Few Labeled Data**
**LIST OF ABSTRACTS**

**Monday July 27th 2020**

---

**Langevin Dynamics : Old and News**

*Eric Moulines*

**Abstract:**
In this keynote, we study a method to sample from a target distribution having a positive density with respect to the Lebesgue measure, known up to a normalization factor. This method is based on the Euler discretization of the overdamped Langevin stochastic differential equation associated with the target distribution. For both constant and decreasing step sizes in the Euler discretization, we obtain non-asymptotic bounds for the convergence to the target distribution in Wasserstein and total variation distance. A particular attention is paid to the dependency on the dimension $d$, to demonstrate the applicability of this method in the high dimensional setting.

**References:**

---

**Learning Physics from Data**

*Francisco Chinesta*

**Abstract:**
Acquiring knowledge from data can be performed in a supervised or an unsupervised way. Particular and still open difficulties concern the data themselves: useful, useless, …; their completeness with respect to the phenomena that we are trying to model (explain), the discovering of the particular form that variables combine to act on the targeted output. Modelling in form of more or less complex regressions can be performed from panoply of techniques, however, in physics first principles must be preserved, fact the enforce constraints but at the same time reduces the amount of need data to perform the learning. Finally, learning will be thermodynamically approached.

**References:**
Information Geometry and Integrable Systems
Jean-Pierre Françoise

Abstract:
We analyze in parallel the (open) Toda-Lattice and the finite Peakons\anti-Peakons system. Their scattering theory relies on a theorem of Stieljes as shown by J. Moser (1975) and R. Beals; D. Sattinger; J. Szmigielski (1999, 01,05,07). We show that both these systems linearize in the setting of Information Geometry. This can be seen as revisiting of previous works of Nakamura, Nakamura and Kodama (1994-1995) where the tau-function of the Toda-Lattice was discovered using Information Geometry.

References:

Schroedinger’s problem, HJB equations
Jean-Claude Zambrini

Abstract:
In 1931-2 Schrödinger formulated an unorthodox problem of classical statistical physics, motivated by his worries about the interpretation of quantum theory. The framework founded on
the solution of his problem is a curious anticipation of Feynman's path integral approach (but probabilistically sound) where Hamilton-Jacobi-Bellman equations are central. We shall describe the connections between these ideas. And why Schrödinger's problem is regarded today as a regularized Monge – Kantorovich problem, at the foundation of Mass transportation theory.

References:

The Bracket Geometry of Measure-Preserving Flows and Diffusions

Alessandro Barp

Abstract:
Following ideas from Koszul, de Rham, and Weinstein, we discuss the canonical geometry generated by a target measure [1-2], and derive characterisations of measure-preserving flows that allows us to extend the complete recipe of stochastic gradient MCMC to manifolds [3].

References:

Sampling and statistical physics via symmetry

Steve Huntsman

Abstract:
In the first part of the talk, we describe how elementary considerations of symmetry (viz., the Lie group preserving a probability measure) lead to a unifying picture of Markov chain Monte Carlo algorithms, including an apparently new parallel MCMC algorithm that converges faster than state-of-the-art techniques. In the second part of the talk, we use basic physical symmetries to parsimoniously derive an effective temperature for steady-state systems with finitely many states. We then show how this construction can be adapted to archetypal physical systems (viz.,
Anosov flows) and produce results suggesting how it may ultimately be used to recover physics from data as well as for more conceptually straightforward descriptive tasks.

References:
**Geometric Mechanics:**

**Gallilean Mechanics & Thermodynamics of Continua**

Géry de Saxcé

Abstract:

Inspired from the relativistic approaches by Souriau [1] and Vallée [2], we propose a geometrization of the thermodynamics of dissipative continua compatible with the Galilean physics [3]. With this aim in view, we emphasize the role of Bargmann’s group [4], a central extension of Galileo’s one [5]. Originally introduced for applications to quantum mechanics, it turns out to be also very useful in thermodynamics.

References:


---

**Geometric Mechanics:**

**Souriau-Casimir Lie Groups Thermodynamics & Machine Learning**

Frédéric Barbaresco

Abstract:

50 years ago, Jean-Marie Souriau invented a “Lie Groups Thermodynamics” model in Statistical Mechanics in his book on Geometric Mechanics, entitled “Structure des systèmes dynamiques” ([http://www.jmsouriau.com/structure_des_systemes_dynamiques.htm](http://www.jmsouriau.com/structure_des_systemes_dynamiques.htm)) and in [1]. We will extend this model in the framework of Information Geometry for Lie Group Statistics and Lie Group Machine Learning. Based on this Symplectic model of Statistical Physics, we can define a Souriau-Koszul-Fisher Metric elaborated on covariant Souriau Gibbs density. Finally, we will propose a new geometric definition of Entropy as a generalized Casimir invariant function in coadjoint representation where Souriau cocycle is a measure of the lack of equivariance of the moment mapping. Lie algebra cohomology, coadjoint orbit methods, and affine representation of Lie Group and Lie Algebra are the main structures used for this elaboration.

References:

Abstract: Inspired by Souriau’s symplectic generalization of Gibbs equilibrium in Lie group thermodynamics, we define a general-covariant notion of Gibbs state for parametrised field theories, in terms of the covariant momentum map associated with the lifted action of the diffeomorphisms group on the extended multi-symplectic phase space. The equilibrium entangles gauge and dynamic information carried by the theory. We investigate how physical equilibrium, hence time evolution, emerges from such a state and the role of the gauge symmetry in the thermodynamic description.

References:

Abstract:
The framework of classical Mechanics is a finite-dimensional Riemannian manifold. Information Geometry describes statistical models as manifolds of probability distributions, endowed with a Riemannian metric given by the Fisher matrix and with a dually flat connection. In this talk we study the mechanic of the probability simplex using an approach based on non parametric
Information Geometry. We start by defining the statistical bundle as the set of couples of a positive probability density and a random variable centered in the density, and we express densities in the affine atlases of exponential charts. We compute velocities and accelerations of a one-dimensional statistical model using the canonical dual pair of parallel transports and define Lagrangian and Hamiltonian mechanics on the bundle. Following our mechanical approach, we are able to define a coherent theory of second-order differential equations which can be used to define different accelerated natural gradient dynamics on the probability simplex.

References:

Diffeological Fisher Metric
Hong Vân Lê

Abstract:
Diffeological Fisher metric is a natural generalization of the Fisher metric on parametrized statistical models to the case of diffeological statistical models, which are statistical models endowed with a compatible diffeology. In my lecture, I shall discuss properties of the diffeological Fisher metric, in particular a diffeological version of the Cramér-Rao inequality.

References:

Contact geometry and thermodynamical systems
Manuel de Leon

Abstract:
Using contact geometry we give a new characterization of a simple but important class of thermodynamical systems which naturally satisfy the first law of thermodynamics (total energy preservation) and the second law (increase of entropy). We clarify its qualitative dynamics, the underlying geometrical structures and we show how to use discrete gradient methods.

References:

Deep learning as optimal control and structure preserving deep learning
Elena Celledoni

Abstract:
Over the past few years, deep learning has risen to the foreground as a topic of massive interest, mainly as a result of successes obtained in solving large-scale image processing tasks. There are multiple challenging mathematical problems involved in applying deep learning. We consider recent work of Haber and Ruthotto 2017 and Chang et al. 2018, where deep learning neural networks have been interpreted as discretisations of an optimal control problem subject to an ordinary differential equation constraint. We review the first order conditions for optimality, and the conditions ensuring optimality after discretisation. There is a growing effort to mathematically understand the structure in existing deep learning methods and to systematically design new deep learning methods to preserve certain types of structure in deep learning. Examples are invertibility, orthogonality constraints, or group equivariance, and new algorithmic frameworks based on conformal Hamiltonian systems and Riemannian manifolds.

References:
« Structure des systèmes dynamiques » [1], now translated into English [2], is a work with an exceptional wealth which, fifty years after its publication, is still topical. We shall intend to highlight author’s most creative and promising ideas on the symplectic geometry and its applications: both classical and relativistic mechanics, geometric quantization and Lie group thermodynamics.

References:
**Thermodynamic efficiency implies predictive inference**

*Susanne Still*

Abstract:

Machine learning is a core ingredient of contemporary statistical data analysis. As with statistics, the foundations are mathematical in nature, often based on “ad hoc” measures. But learning is a natural phenomenon occurring in the physical world. Therefore, we would like to have a physics based explanation. We need to understand how observers choose their strategy for how to represent, and adapt to, the data they receive. Indulge, for a moment, the following hypothesis: observers choose their strategy such that the best physical implementation of abstract rules specifying the strategy could come as close as possible to the physical limits imposed on information processing. This postulate would open a door to “derivative” learning methods from “physics”, simply by minimizing a physical bound over all possible rules, thereby finding the strategy optimal with respect to the limitation expressed by the bound. Here, I show that this can be done for thermodynamic limits: energy efficiency implies predictive inference, a strategy that lies at the heart of machine learning.

References:


---

**Exponential Family by Representation Theory**

*Koichi Tojo*

Abstract:

Exponential families play an important role in the field of information geometry. By definition, there are infinitely many exponential families. However, only a small part of them are widely used. We want to give a framework to deal with these "good" families. In light of the observation that the sample space of most of them are homogeneous spaces of certain Lie groups, we proposed a method to construct exponential families on homogeneous spaces by taking advantage of representation theory in [1]. This method generates widely used exponential families such as normal, gamma, Bernoulli, categorical, Wishart, von Mises-Fisher, and hyperboloid distributions. In this talk, we will explain the method, its properties and future works.

References:

<table>
<thead>
<tr>
<th>Poster Authors</th>
<th>Poster Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timothee Pouchon, Benedict Leimkuhler, Charles Matthews and Tiffany Vlaar</td>
<td>Constraint-Based Regularization of Neural Networks</td>
</tr>
<tr>
<td>Kevin Grosvenor</td>
<td>Information Geometry and the Effective Field Theory of Flocking</td>
</tr>
<tr>
<td>Rita Fioresi</td>
<td>A geometric interpretation of stochastic gradient descent in Deep Learning and Restricted Boltzmann Machines</td>
</tr>
<tr>
<td>Filipe Dias</td>
<td>Geometric Thermodynamics of Information Processing and Fluctuations</td>
</tr>
<tr>
<td>Anis Fradi and Chafik Samir</td>
<td>Bayesian Inference on Local Distributions of Functions and Multi-dimensional Curves with Spherical HMC Sampling</td>
</tr>
<tr>
<td>Carlos Couto, José Mourão, João P. Nunes and Pedro Ribeiro</td>
<td>Connecting Stochastic Optimization with Schrödinger evolution with respect to non-Hermitian Hamiltonians</td>
</tr>
<tr>
<td>Emmanuel Chevallier and Nicolas Guigui</td>
<td>Warped statistical models on $SE(n)$: motivation, challenges and generalization on symmetric spaces</td>
</tr>
<tr>
<td>Sébastien Boyaval</td>
<td>Viscoelastic flows with conservation laws</td>
</tr>
<tr>
<td>Filippo Masi, Ioannis Stefanou, Paolo Vannucci and Victor Maffi-Berthier</td>
<td>Material modeling via Thermodynamics-based Artificial Neural Networks</td>
</tr>
<tr>
<td>Nicolas Guigui, Nina Miolane and Alice le Brigant</td>
<td>Geomstats: a Python package for Geometric Learning and Information Geometry</td>
</tr>
<tr>
<td>Hatem Hajri, Thomas Gerald and Hadi Zaatiti</td>
<td>Hyperbolic learning of communities on graphs</td>
</tr>
<tr>
<td>Goffredo Chirco, Luigi Malagò and Giovanni Pistone</td>
<td>Lagrangian and Hamiltonian Dynamics on the Simplex</td>
</tr>
<tr>
<td>Héctor Javier Hortúa, Luigi Malagò and Riccardo Volpi</td>
<td>Calibrating Bayesian Neural Networks with Alpha-divergences and Normalizing Flows</td>
</tr>
<tr>
<td>Youssef El Habouz</td>
<td>Semi-supervised Classification of Cells based on GAN</td>
</tr>
<tr>
<td>Avetik Karagulyan and Arnak Dalalyan</td>
<td>Bounding the error of discretized Langevin algorithms for non-strongly log-concave targets</td>
</tr>
<tr>
<td>Elvis Dohmatob</td>
<td>Universal Lower-Bounds on Classification Error under Adversarial Attacks and Random Corruption</td>
</tr>
<tr>
<td>Pierre-Cyril Aubin-Frankowski and Zoltan Szabo</td>
<td>Hard Shape-Constrained Kernel Regression</td>
</tr>
<tr>
<td>Bruno Sauvalle</td>
<td>Unsupervised object detection for traffic scene analysis</td>
</tr>
<tr>
<td>Paul Ferrand, Alexis Decurninge, Luis Garcia Ordonez and Maxime Guillaud</td>
<td>Learning the low-dimensional geometry of the wireless channel</td>
</tr>
</tbody>
</table>
Non-Equilibrium Thermodynamic Geometry: A variational perspective on nonequilibrium thermodynamics of closed and open systems
François Gay-Balmaz

Abstract:
We survey recent results on the variational formulation of nonequilibrium thermodynamics for finite-dimensional and continuum systems. We illustrate the theory with closed and open systems experiencing friction, heat and mass transfer, and chemical reactions. We show how the theory is used for discretization and as a modeling tool in fluid dynamics.

References:

Non-Equilibrium Thermodynamic Geometry: A Homogeneous Symplectic Approach
Arjan van der Schaft

Abstract:
Since the early 1970s contact geometry has been recognized as an appropriate geometric framework for thermodynamic systems. In the 2001 paper by Balian and Valentin it was shown how the homogeneous symplectic approach to contact geometry has several advantages, e.g., in switching between energy and entropy representations. In this talk I will show how this approach leads to a geometric formulation of non-equilibrium thermodynamic processes, in terms of Hamiltonian dynamics defined by Hamiltonian functions that are homogeneous of degree one in the co-extensive variables and zero on the homogeneous Lagrangian submanifold describing the state properties. This culminates in the definition of port-thermodynamic systems, and the formulation of interconnection ports with the environment or other systems. This is illustrated on a number of simple examples, indicating its potential for analysis and control.

References:
Dirac structures and port-Dirac systems in nonequilibrium thermodynamics

Hiroaki Yoshimura

Abstract:
A Dirac structure is a unifying notion of symplectic and Poisson structures, which has been widely used in mechanics, in particular, for mechanical systems with nonholonomic constraints. In this talk, we study Dirac structures in nonequilibrium thermodynamics by extending to a class of nonlinear nonholonomic systems. We also clarify the associated variational structures together with some examples of open systems as well as interconnected systems. This is a joint work with Francois Gay-Balmaz.

References:

Port Thermodynamic Systems Control

Bernhard Masche

Abstract:
We consider the feedback control of homogeneous Hamiltonian control systems arising in the Hamiltonian modelling of open thermodynamic systems and presented in the talk “Non-Equilibrium Thermodynamic Geometry: A Homogeneous Symplectic Approach”. In this talk we characterize classes of state feedbacks for which the closed-loop system is again Homogeneous Hamiltonian and leaves invariant some closed-loop 1-form and derive some relations with the closed-loop Hamiltonian function.

References:
Computational dynamics of reduced coupled multibody-fluid system
in Lie group setting

Zdravko Terze

Abstract:
We describe a computationally efficient method for simulating dynamics of multibody-fluid system that utilizes symplectic and Lie-Poisson reductions in order to formulate fully coupled dynamical model of the multi-physical system by using solid variables only. Multibody system dynamics is formulated in Lie group setting and integrated with the pertinent integration method, while additional viscous effects are incorporated in the overall model by numerically enforcing Kutta condition.

References:

Information Geometry & Quantum Fields

Kevin Grosvenor

Abstract:
We study the Fisher metrics associated with a variety of simple systems and derive some general lessons that may have important implications for the application of information geometry in holography. Some sample systems of interest are the classical 2d Ising model and the corresponding 1d free fermion theory, massless scalar instantons, and coherent states of free bosons and fermions.

References:

Learning with Few Labeled Data
Pratik Chaudhari

Abstract:
The relevant limit for machine learning is not $N \to \infty$ but instead $N \to 0$, the challenge is to build systems that do not require $N = \text{thousands of labeled data}$. We will exploit a formal connection of thermodynamics and machine learning to characterize the limits of representation learning in the low-data regime. This theory leads to algorithms that can guarantee good classification performance after the model is transferred onto a new task.

References:
**Computational Information Geometry:**

On statistical distances and information geometry for machine learning

*Frank Nielsen*

**Abstract:**

We survey recent progress in the construction of divergences and their induced information geometry with applications to machine learning: First, we provide generalizations of the celebrated Jensen-Shannon divergence [1,2] that is at the heart of Generative Adversarial Networks. Second, we describe some statistical divergences on the Cauchy manifold [3] with their information-geometric structures, and show applications in statistics. Third, we show how to quickly calculate numerically the Siegel distance on the Siegel disk domain using the novel Siegel-Klein model [4] based on Hilbert geometry, and discuss applications in machine learning. Last, we show a simple trick to easily calculate statistical distances between exponential families using legacy software packages [5].

**References:**


**Coffe Break**

**Computational Non-Parametric Information Geometry**

Information Manifold modeled with Orlicz Spaces

*Giovanni Pistone*

**Abstract:**

One of the possible non-parametric version of Information Geometry with infinite sample space assumes strictly positive densities whose logarithm belongs to an Orlicz space. In this way, the full structure of the affine Hessian statistical manifold can be rigorously derived. After a brief summary of this old theory, I will discuss some recent developments:

1. The extension of the Orlicz space modeling to the statistical bundle;
2. The use of Orlicz-Sobolev spaces to allow for smoothness of the densities;
3. The special features of the finite-dimensional Gaussian Space.

**References:**
