Abstracts

Frédéric Barbaresco
Thales Group, France

Symplectic Foliation-informed Neural Network (SFINN) and Lie Groups Machine Learning based on Jean-marie Souriau Lie Groups Thermodynamics & Koszul Information Geometry

The symplectic model of statistical mechanics developed by Jean-Marie Souriau, termed the "Thermodynamics of Lie Groups", extends the structures of information geometry to the realm of Lie groups. This framework enables the definition of Maximum Entropy Gibbs densities possessing the property of covariance under the action of the group operating on the system. Moreover, it generalises the Fisher-Rao-Fréchet metric to Lie groups, rendering it invariant under the group's action. Crucially, Shannon's axiomatic definition of entropy is supplanted by a purely geometric construction, wherein entropy emerges as a Casimir invariant function defined on the leaves of the foliation induced by coadjoint orbits through the moment map associated with the group action (the moment map being the geometric counterpart of Noether's theorem).

Souriau's thermodynamics of Lie groups introduces a web-like geometric structure composed of two transverse foliations: a symplectic foliation generated by coadjoint orbits (corresponding to the level sets of entropy) and a transverse Riemannian foliation (corresponding to the level sets of energy). The dynamics on each foliation make it possible to distinguish between non-dissipative phenomena (with constant entropy) and dissipative phenomena (with constant energy and entropy production). This dynamic behaviour is governed by a metriplectic flow that encapsulates the first law of thermodynamics through Poisson bracket (quantitative conservation of energy) and the second law through metric flow bracket (qualitative degradation of energy and generation of entropy).

We shall explore the connections between TINNs (Thermodynamics-Informed Neural Networks), metriplectic flows, and the Lie groups thermodynamics. The overarching aim is for TINNs not merely to learn from data, but also to adhere to thermodynamic constraints, thereby enabling more accurate predictions and a deeper understanding of physical systems, particularly those characterised by dissipative phenomena.

Souriau Lie Groups Thermodynamics is studied in the framework of two European action, European CaLISTA COST action and European CaLIGOLA MSCA action.

REFERENCES

[1] Barbaresco, F.; Gay-Balmaz, F. (2020) Lie Group Cohomology and (Multi)Symplectic Integrators: New Geometric Tools for Lie Group Machine Learning Based on Souriau Geometric Statistical Mechanics. Entropy, 22, 498

[2] Barbaresco F (2021) Koszul lecture related to geometric and analytic mechanics, Souriau's Lie group thermodynamics and information geometry. Information Geometry Journal, Springer 4:245–262.

[3] Barbaresco F (2021) Jean-Marie Souriau's Symplectic Model of Statistical Physics: Seminal Papers on Lie Groups Thermodynamics - Quod Erat Demonstrandum. In: Nielsen F and Barbaresco F (eds). Geometric Structures of Statistical Physics, Information Geometry, and Learning, Springer, Berlin, pp. 12–50.

[4] Barbaresco F (2022) Symplectic theory of heat and information geometry. Handbook of Statistics, special issue Geometry and Statistics 46(4):107-143.

[5] Barbaresco F (2022) Symplectic Foliation Structures of Non-Equilibrium Thermodynamics as Dissipation Model: Application to Metriplectic Nonlinear Lindblad Quantum Master Equation. Entropy 24 :1626-1662.

[6] Barbaresco F (2022) Symplectic foliation model of information geometry for statistics and learning on Lie groups. Geometry, topology and statistics in data sciences, Institut Henri Poincaré, Paris, https://www.carmin.tv/fr/video/symplectic-foliation-model-of-information-geometry-for-statistics-and-learning-on-lie-groups

[7] Barbaresco F (2023) Symplectic Foliation Transverse Structure and Libermann Foliation of Heat Theory and Information Geometry. In: Nielsen F and Barbaresco F (eds) Geometric Science of Information GSI'23. Lecture Notes in Computer Science 14072, Springer, Berlin, pp. 152–164.

[8] Barbaresco F (2025), Transverse Symplectic Foliation Structure for Thermodynamics-Informed Neural Network and Lie-Groups Machine Learning, Erwin Schrödinger Institute Seminar, Infinite-dimensional Geometry: Theory and Applications trimester, https://www.esi.ac.at/events/t2213/

[9] Barbaresco, F. (2025) Jean-Marie Souriau's Symplectic Foliation Model of Sadi Carnot's Thermodynamics. Entropy, 27, 509. https://doi.org/10.3390/e27050509

[10] Barbaresco, F. (2025) Transverse Symplectic Foliation Structure for Dissipative Thermodynamics based on Souriau Model of Statistical Mechanics, 18th Joint European Thermodynamics Conference, Mini-Symposia "Foundations and philosophy of thermodynamics", Belgrade

[11] Souriau JM (1969) Structure des systems Dynamique, Dunod, translated in English "Structure of Dynamical Systems: A Symplectic View of Physics" Birkhauser 1912, https://empg.maths.ed.ac.uk/Activities/Souriau/Souriau.pdf

[12] Souriau JM (1974) Mécanique statistique, groupes de Lie et cosmologie. In : Souriau JM (eds). Colloque International du CNRS Géométrie symplectique et physique Mathématique, CNRS, Marseille.

[13] Koszul JL, Zou YM (2019), Introduction to Symplectic Geometry, Springer Nature Singapore, https://link.springer.com/book/10.1007/978-981-13-3987-5

[14] Barbaresco F, Nielsen Frank (2021) Geometric Structures of Statistical Physics, Information Geometry, and Learning: SPIGL'20, Les Houches, France, July 27–31, Springer, https://link.springer.com/book/10.1007/978-3-030-77957-3

[15] European COST network CaLISTA: https://site.unibo.it/calista/en

[16] European HORIZON-MSCA CaLIGOLA https://site.unibo.it/caligola/en

[17] Special Issue, Dynamics Beyond the Hamiltonian: Dissipation in Classical Metriplectic Systems and Quantum Non-Unitary Systems, MDPI Entropy, Guest Editors Dr. Massimo Materassi & Frédéric Barbaresco, mdpi.com/si/229747

Tolga Birdal
Imperial College London, UK

Topological Complexity Measures as Proxies for Generalization in Neural Networks

Deep neural networks (DNNs) exhibit remarkable generalization abilities, yet the mechanisms behind these capabilities remain poorly understood, defying the established wisdom of statistical learning theory. Recent research has revealed a compelling link between the fractal structures formed during iterative training and the resulting generalization performance. In this talk, Dr. Birdal sheds new light on these connections by presenting a novel framework that ties complexity measures to the topological properties of the training process.

The presentation begins by bounding the generalization error through the fractal dimension of training trajectories, practically computed using tools from persistent homology—introducing the 'persistent homology dimension' as a new, insightful proxy for generalization. Building on this, Dr. Birdal introduces more computationally efficient topological complexity measures that bypass the need for continuous training trajectories. These measures consistently show strong correlations with the generalization gap across diverse models, including transformers and graph networks. The findings hold transformative implications for both theory and practice, offering a new lens to study, understand and optimize the generalization power of modern AI systems.

Relevant Publications:

[1] https://arxiv.org/abs/2111.13171 [arxiv.org] (NeurIPS 2021)
[2] https://arxiv.org/abs/2407.08723 [arxiv.org] (NeurIPS 2024)

Back to Contents Page

Cristian Bodnar
Silurian AI, UK

Aurora: A Foundation Model for the Earth System

Reliable forecasts of the Earth system are crucial for human progress and safety from natural disasters. Artificial intelligence offers substantial potential to improve prediction accuracy and computational efficiency in this field, however this remains underexplored in many domains. Here we introduce Aurora, a large-scale foundation model for the Earth system trained on over a million hours of diverse data. Aurora outperforms operational forecasts for air quality, ocean waves, tropical cyclone tracks, and high-resolution weather forecasting at orders of magnitude smaller computational expense than dedicated existing systems. With the ability to fine-tune Aurora to diverse application domains at only modest computational cost, Aurora represents significant progress in making actionable Earth system predictions accessible to anyone.

Back to Contents Page

Baris Coskunuzer
University of Texas Dallas, USA

Topological Compound Fingerprinting in Computer-Aided Drug Discovery

In this talk, I present a novel topology-based approach to virtual screening in computer-aided drug discovery using multiparameter persistence. Traditional methods relying on SMILES strings, molecular fingerprints, or deep learning models like VAEs and GNNs face challenges in scalability and performance saturation. Our method generates multidimensional topological fingerprints by decomposing compounds into chemically informed substructures and extracting persistent homology features at multiple resolutions. We reformulate VS as a graph ranking problem and use few-shot learning techniques to effectively rank compounds by drug-likeness. Our approach achieves substantial performance gains over state-of-the-art methods on benchmark datasets. For more details, please refer to our paper: https://arxiv.org/abs/2211.03808

Back to Contents Page

Xiaowen Dong
University of Oxford, UK

Bayesian Optimisation of Graph-based Functions

The increasing availability of graph-structured data motivates a new type of optimisation problems over graph-based functions, i.e., searching for the graph or node that maximises the value of an underlying function. Such optimisation problems are challenging due to the discrete and high-dimensional search space, as well as the underlying function that is often black-box and expensive to evaluate. In this talk, I will provide several examples on how Bayesian optimisation can be used to optimise graph-based functions, with practical applications in computational, epidemiological, and social networks. More broadly, these examples demonstrate the promise in combining probabilistic and geometric reasoning in analysing complex functions.

Back to Contents Page

Mustafa Hajij
University of San Francisco, USA

Frontiers and Opportunities in Topological Deep Learning

In recent years, deep learning has achieved remarkable success across various domains, particularly in tasks involving data structured as regular grids or sequences, such as images and text. However, scientific and real-world data frequently exhibit complex, non-Euclidean structures—including point clouds, meshes, graphs, and higher-order topological spaces—that challenge the assumptions of traditional neural network architectures. Topological Deep Learning (TDL) is an emerging field that extends deep learning methods to handle these rich and intricate data types. By incorporating topological constructs such as simplicial complexes, cell complexes, and hypergraphs, TDL enables the modelling of higher-order relationships, global dependencies, and qualitative spatial properties that are otherwise inaccessible to standard approaches. This talk will explore the foundational ideas behind TDL, the computational and theoretical challenges it presents, and its broad potential to enhance learning in diverse areas—from physics-informed machine learning and neuroscience to social network analysis and beyond.

Back to Contents Page

Niu Huang
National Institute of Biological Sciences, China
Tsinghua Institute of Multidisciplinary Biomedical Research, China

Integrating HPC and AI: A New Paradigm for Predicting Protein-ligand Binding

In the process of small molecule drug discovery, the prediction of protein-ligand interactions urgently demands enhancements in computational accuracy and efficiency, given its crucial role in identifying novel lead compounds for new targets. However, current artificial intelligence (AI) models are constrained by the scarcity of large, high-quality protein-ligand complex structures and binding data, which consequently impairs their generalization ability, limiting their effectiveness in real-world applications. We have been actively exploring the potential of physics-based high performance computing (HPC). The remarkable computational power of HPC allows us to generate vast, top-tier datasets that are invaluable for both training and testing AI modes. When integrated with AI's proficiency in pattern recognition and predictive modelling, this combination allows for the rapid and in-depth analysis of molecular structures, more accurate prediction of drug-target interactions. Our ongoing research and practice will highlight the profound synergy between HPC and AI in facilitating more accurate and efficient calculations of molecular interactions, illuminating viable strategies to surmount existing data limitations and improve the predictive capabilities of AI models.

Back to Contents Page

Wei Huang
RIKEN, Japan

Decoding Deep Graph Neural Networks: An Optimization and Generalization Perspective

Graph Neural Networks (GNNs) have emerged as a powerful framework for modelling graph-structured data, yet challenges such as over-smoothing, trainability degradation in deep architectures, and understanding the role of structural information remain. In this talk, we will explore two complementary research directions addressing these challenges from both optimization and generalization perspectives.

First, we will delve into the optimization dynamics of deep GNNs using a Graph Neural Tangent Kernel (GNTK) framework. This approach reveals an exponential decline in trainability as network depth increases—a phenomenon that standard residual connections can only partially mitigate. Based on these insights, we introduce the Critical DropEdge method, a connectivity-aware and graph-adaptive sampling strategy designed to fundamentally counteract this decay and enhance model performance.

In the second part of the talk, the focus shifts to the feature learning capabilities of GNNs. We will demonstrate how graph convolution, by leveraging inherent structural information, significantly amplifies signal learning while suppressing noise memorization. Comparative analyses with multilayer perceptrons (MLPs) highlight that graph convolution offers superior generalization.

Back to Contents Page

Stephan Huckemann
Georg-August-Universität Göttingen, Germany

Dirty Limit Theorems and Applications

Fréchet means are generalizations of the expected value to manifolds and stratified spaces. Their asymptotic rates, however, may deviate from those of their Euclidean kin. For instance, faster rates let the sample mean "stick" to the population mean, whereas slower rates let it appear to be "smearily" spread out. In this talk we illustrate some relationships between geometry and statistics via dirty (sticky, smeary) asymptotic rates of Fréchet means. We illustrate the relevance of these results for statistical testing for geometrical shape.

Back to Contents Page

Stephan Klaus
Mathematisches Forschungsinstitut Oberwolfach, Germany

Nonlinear Regression with Real Algebraic Varieties and their Topology

The standard approach in topological data analysis is by persistent homology using the Čech or Vietoris-Rips complex for a point data cloud. This approach has a problem of high computational cost.
In our talk we consider another possible method: approximation of the point cloud by nonlinear regression with real algebraic varieties. In addition, we give an overview on some results concerning the topology of real algebraic varieties.

Back to Contents Page

Patrice Koehl
University of California, Davis, USA

A Physicist's View on Partial 3D Shape Comparison

Scientists have access to a wide range of digital sensors that allow them to report at multiple time and length scales on the subjects of their studies. Finding efficient algorithms to describe and compare the shapes included in those reports has become a central problem in data science. Those algorithms have gained from developments in computational geometry and in machine learning. In this talk I will present another source of support to further improve those algorithms. Using techniques from statistical physics, I show that we can define a possibly partial correspondence between 3D shapes, with a cost associated with this correspondence that serves as a measure of the similarity of the shapes. I will illustrate the effectiveness of this approach on synthetic data as well as on real anatomical data.

This is joint work with Dr Henri Orland, IPHT, CEA, Saclay, France.

Back to Contents Page

Ran Levi
The University of Aberdeen, UK

Foundations of Differential Calculus for Modules over Small Categories

Let k be a field and let C be a small category. A k-linear representation of C, or a kC-module, is a functor from C to the category of finite dimensional vector spaces over k. A motivating example for this work is the concept of a tame generalised persistence module, which can be reduced to the case where C is a finite poset. Unsurprisingly, it turns out that when the category C is more general than a linear order, then its representation type is generally infinite and in most cases wild. Hence the task of understanding such representations in terms of their indecomposable factors becomes difficult at best, and impossible in general. In a joint project with Jacek Brodzki and Henri Rihiimaki we proposed a new set of ideas designed to enable studying modules locally. Specifically, inspired by work in discrete calculus on graphs, we set the foundations for a calculus type analysis of kC-modules, under some restrictions on the category C. In this talk I will review the basics of the theory and describe some more recent advances.

Back to Contents Page

Zheng Ma
Shanghai Jiao Tong University, China

Solving PDE Inverse Problems with Generative Models and Their Applications

While deep learning has advanced PDE inverse problem solutions, current methods often depend on paired data or retraining when conditions change—limiting efficiency and flexibility. To overcome these challenges, we present an unsupervised inversion framework leveraging score-based generative diffusion models within a Bayesian inversion paradigm. Our approach recasts posterior estimation as a conditional generative process using reverse-time stochastic differential equations (SDEs). Furthermore, we introduce a diffusion posterior sampling algorithm based on ordinary differential equations (ODEs), ensuring accuracy through marginal probability consistency across forward Fokker-Planck dynamics. Experiments validate robust performance across diverse PDEs.

Back to Contents Page

Frank Nielsen
Sony Computer Science Laboratories, Japan

Computational Information Geometry on Bregman Manifolds and Submanifolds

We review the construction of a Bregman manifold from a Legendre-type convex function [1]. By further using a representation function, we show that alpha-divergences are representational Bregman divergences on the positive orthant cone and curved representational Bregman divergences on the probability simplex [2]. We describe clustering [3], nearest-neighbour query data structures [4], and Voronoi diagrams [5] on Bregman manifolds and submanifolds with several applications in statistics and data science. Finally, we present work in progress pyBregMan: A Python library for algorithms and data-structures on BREGman MANifolds [6].

References:

[1] "On geodesic triangles with right angles in a dually flat space." Progress in Information Geometry: Theory and Applications. Cham: Springer International Publishing, 2021. 153-190.
[2] "The dual Voronoi diagrams with respect to representational Bregman divergences." 2009 Sixth International Symposium on Voronoi Diagrams. IEEE, 2009.
[3] "On clustering histograms with k-means by using mixed α-divergences." Entropy 16.6 (2014): 3273-3301.
[4] "Bregman vantage point trees for efficient nearest neighbor queries." 2009 IEEE International Conference on Multimedia and Expo. IEEE, 2009.
[5] "Bregman voronoi diagrams." Discrete & Computational Geometry 44 (2010): 281-307.
[6] PyBregMan: https://franknielsen.github.io/pyBregMan/

Back to Contents Page

Anthea Monod
Imperial College London, UK

Algebraic Geometry Learns Machines and
Machines Learn Algebraic Geometry

In this talk I will overview some existing results and ongoing work at the intersection of algebraic geometry and machine learning. I will present how a piecewise linear and combinatorial variant of algebraic geometry—-known as tropical geometry—-has been shown to be relevant in defining neural networks and talk about some recent and current work that our group is doing that adapts tropical geometry theory in numerical studies towards a better understanding of neural network behaviour during training. While algebraic geometry holds much potential for better understanding machine learning, it turns out that machine learning is also a powerful tool that can help develop algebraic geometry theory. I will also overview some recent and ongoing work by researchers in my group where we use neural networks for theorem discovery in algebraic geometry.

Back to Contents Page

Hans Riess
Georgia Institute of Technology, USA

Categories and Sheaves for Optimization:
From Multi-Stage to Distributed

Optimization techniques are central to data science, from training neural networks to collaborative filtering to federated learning. In this talk, we argue category theory and sheaf theory enhance optimization techniques. We first show that multi-stage convex and non-convex optimization problems can be modelled with enriched category theory. Motivated by coordination problems over networks, we then discuss distributed optimization through the lens of homological programming and introduce sheaf Laplacians to help solve homological programs. To conclude the talk, we present recent advances in sheaf Laplacians and further promising directions.

Back to Contents Page

Roman Sauer
Karlsruhe Institute for Technology, Germany

Expanders, Waists, and the Kazhdan Property

The Kazhdan property, introduced by Kazhdan in the 60s, is a property of unitary representations of a group. Margulis used the Kazhdan property to provide the first explicit construction of expander graphs which found numerous applications in computer science. In this talk we discuss a Riemannian analog of expander graphs as well as higher-dimensional generalizations. This based on joint work with Uri Bader.

Back to Contents Page

Jian Tang
HEC Montréal, Canada

Geometric Deep Learning for Protein Design

Proteins are workhorses of living cells. Understanding the functions of proteins is critical to many applications such as biomedicine and synthetic biology. Thanks to recent biotechnology breakthroughs such as gene sequencing and Cro-EM, a large amount of protein data (such as protein sequences and structures) are generated, providing a huge opportunity for AI. As the functions of proteins are determined by their structures, in this talk, I will introduce our recent work on protein understanding based on protein 3D structures with geometric deep learning. I will introduce three different topics including protein representation learning, generative models for protein structure prediction, and generative models for protein design, and also how these techniques are used for real-world applications in protein design.

Back to Contents Page

Ben Yang
University of Oxford, UK
Max Planck Institute of Molecular Cell Biology and Genetics, Germany

Large Isometry Invariant Topological Transform Shape Descriptor

The Euler Characteristic Transform (ECT) is a powerful shape descriptor thanks to its invertibility. However, this also means ECT is sensitive to rigid motions, meaning it can distinguish between representations of the same shape. To address this, we introduce two new tools: **SampEuler**, a robust, isometry-invariant shape descriptor based on ECT, and **EulerImage**, a visualization and vectorization of SampEuler. We show that both methods capture sufficient geometric information of the input shape while reducing the effects of rigid motions. We demonstrate their effectiveness on both synthetic and real-world datasets. We also highlight how EulerImage helps interpretations of the result.

Back to Contents Page

Yaoyu Zhang
Shanghai Jiao Tong University, China

Towards Understanding the Condensation Phenomenon of Deep Neural Networks

Condensation (also known as quantization, weight clustering, or alignment) is a widely observed phenomenon where neurons in the same layer tend to align with one another during the nonlinear training of deep neural networks (DNNs). It is a key characteristic of the feature learning process of neural networks. However, due to the strong nonlinear nature of this phenomenon, establishing its theoretical understanding remains challenging. In this talk, I will present our systematic efforts to tackle this challenge in recent years. First, I will present results regarding the dynamical regime identification of condensation at the infinite width limit, where small initialization is crucial. Then, I will discuss the mechanism of condensation at the initial training stage and the global loss landscape structure underlying condensation in later training stages, highlighting the prevalence of condensed critical points and global minimizers. Finally, I will present results on the quantification of condensation and its generalization advantage, which includes a novel estimate of sample complexity in the best-possible scenario. These results underscore the effectiveness of the phenomenological approach to understanding DNNs, paving the way for a deeper understanding of deep learning in the near future.

Back to Contents Page

Yipeng Zhang
Nanyang Technological University, Singapore

Multi-Cover: A Mathematical Framework for Topological Data Analysis and Deep Learning

Topological Data Analysis (TDA) has made significant contributions to molecular and materials science. Multi-Cover Persistence (MCP) and its associated Rhomboid Tiling (RT) structure, as a generalization of the alpha shape in TDA, provide a powerful framework for capturing the shape and higher-order geometric features of objects. Leveraging these concepts, we developed three innovative computational models: First, a featurization-based machine learning model where the MCP framework extracts topological features from the persistent homology of molecular multi-covers to predict polymer properties with high accuracy. Second, the RT structure forms the basis for a hierarchical graph pooling model for molecular graph classification tasks. Third, building upon RT's multi-scale hierarchical structure, we developed a topological deep learning model that utilizes RT's higher-order geometric relationships to design an advanced message-passing mechanism. This framework demonstrates particular effectiveness in protein complex quality assessment. All three models show excellent performance, highlighting the versatility of MCP and RT in advancing polymer informatics, geometric graph learning, and topological deep learning applications.

Back to Contents Page

Difan Zou
University of Hong Kong, Hong Kong SAR

Towards Understanding the Representation Learning of Diffusion Models

Diffusion models (DMs) excel in generative modeling, but their theoretical foundations and limitations remain underexplored. This talk addresses two key aspects: their feature learning dynamics and their ability to capture hidden inter-feature rules. First, I show that the denoising objective encourages DMs to learn balanced and comprehensive data representations, unlike classification models that prioritize easy-to-learn patterns. Theoretical analysis and experiments on synthetic and real-world datasets highlight this distinction. Next, I explore a critical limitation: DMs often fail to learn fine-grained hidden rules between dependent features, such as the relationship between the height of the sun and shadow length in images. Empirical evaluations on models like Stable Diffusion reveal consistent failures, supported by synthetic tasks and theoretical insights showing that denoising score matching (DSM) is incompatible with enforcing rule conformity. I discuss potential solutions, such as classifier-guided sampling, and their limitations. This talk provides a deeper understanding of DMs' strengths and weaknesses, offering insights for building more robust and interpretable generative models.

Back to Contents Page