Workshop: Mathematics of data driven models

Monday, January 26, 2026 - Wednesday, January 28, 2026

GSSI

Book of Abstracts

Contents

Session 1 / **3**

Scaling Laws for Hypergraph and Multi-Operator Learning

Author: Adrien Weihs[1]

[1] *University of California Los Angeles*

Corresponding Author: adrien.weihs@icloud.com

This talk presents results on scaling laws in hypergraph and multi-operator learning. Scaling considerations simultaneously guide both the theoretical foundations and the practical design of modern learning systems. The first part shows how a large-data asymptotic analysis identifies connectivity-scaling regimes in which semi-supervised learning on hypergraphs is effective and stable. This perspective also leads to a principled taxonomy of hypergraph learning algorithms, organized by their underlying regularization mechanisms and induced continuum limits. The second part illustrates how scaling insights inform the design of expressive multi-operator networks and provide principled answers to the architectural search problem. Specifically, bounds are derived on the required network width, depth, and sparsity to achieve a prescribed approximation accuracy.

Session 3 / **4**

Latent Dynamics Graph Convolutional Networks for Model Order Reduction

Author: Lorenzo Tomada[1]

Co-authors: Federico Pichi [1]; Gianluigi Rozza [1]

[1] *SISSA*

Corresponding Author: ltomada@sissa.it

Graph Neural Networks (GNNs) have emerged as powerful tools for nonlinear Model Order Reduction (MOR) of time-dependent parameterized Partial Differential Equations (PDEs) [1].
However, existing methodologies struggle to combine geometric inductive biases with interpretable latent dynamics, overlooking dynamics-driven features or disregarding geometric information, respectively.

In this work, we address this gap by introducing Latent Dynamics Graph Convolutional Networks (LD-GCNs) [3], a purely data-driven, encoder-free architecture that learns a global, low-dimensional representation of dynamical systems conditioned on external inputs and/or parameters [2].
The temporal evolution is modeled in the latent space and advanced through time-stepping, allowing for time-extrapolation, and the resulting trajectories are consistently decoded onto geometrically parametrized domains using a GNN.
Our framework enhances interpretability by enabling the analysis of latent trajectories and supports zero-shot prediction through interpolation in the latent space.

The methodology is mathematically validated via a universal approximation theorem for encoder-free architectures, and numerically tested on complex computational mechanics problems involving physical and geometrical parameters, including the detection of bifurcating phenomena for Navier-Stokes equations.

References:
[1] Federico Pichi, Beatriz Moya, and Jan S. Hesthaven. “A graph convolutional autoencoder approach to model order reduction for parametrized PDEs”. In: *Journal of Computational Physics* 501 (Mar. 2024), p. 112762. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2024.112762.
[2] Francesco Regazzoni et al. “Learning the intrinsic dynamics of spatio-temporal processes through Latent Dynamics Networks”. en. In: *Nature Communications* 15.1 (Feb. 2024), p. 1834. ISSN: 2041-1723. DOI:

10.1038/s41467-024-45323-x
[3] Lorenzo Tomada, Federico Pichi, and Gianluigi Rozza. *Latent Dynamics Graph Convolutional Networks for Model Order Reduction.* In preparation.

Session 5 / **5**

Gradient flow-based modularity maximization for community detection in multiplex networks

Author: Kai Bergermann[1]

Co-author: Martin Stoll [2]

[1] *Scuola Normale Superiore, Pisa*

[2] *TU Chemnitz*

Corresponding Author: kai.bergermann@sns.it

We propose two methods for the unsupervised detection of communities in undirected multiplex networks. These networks consist of multiple layers that record different relationships between the same entities or incorporate data from different sources. Both methods are formulated as gradient flows of suitable energy functionals: the first (MPBTV) builds on the minimization of a balanced total variation functional, which we show to be equivalent to multiplex modularity maximization, while the second (DGFM3) directly maximizes multiplex modularity. The resulting non-linear matrix-valued ordinary differential equations (ODEs) are solved efficiently by a graph Merriman–Bence–Osher (MBO) scheme. Key to the efficiency is the approximate integration of the discrete linear differential operators by truncated eigendecompositions in the matrix exponential function. Numerical experiments on several real-world multiplex networks show that our methods are competitive with the state of the art with respect to various metrics. Their major benefit is a significant reduction of computational complexity leading to runtimes that are orders of magnitude faster for large multiplex networks.

Session 4 / **6**

Nearest correlation matrices with structure: a dynamical systems approach

Authors: Nicola Guglielmi[1]; Christian Lubich[2]; Francesco Paolo Maiale[3]; Eugenio Turchet[1]

[1] *Gran Sasso Science Institute*

[2] *University of T¨ubingen*

[3] *GSSI*

Corresponding Author: eugenio.turchet@gssi.it

The nearest correlation matrix problem consists in finding the closest valid correlation matrix to a given symmetric matrix that may fail to be positive semi-definite. In other words, given a symmetric unit-diagonal matrix that is not a proper correlation matrix, one seeks the nearest positive semi-definite matrix with unit diagonal entries.

We address the problem of finding the nearest correlation matrix to a given symmetric unit-diagonal matrix under additional structural constraints such as sparsity, block, or band patterns. This task arises in applications where positive semi-definiteness must be restored without losing essential structure.

Our method combines a two-level iteration: a structured gradient flow computes feasible perturbations within the prescribed structure, while an outer Newton scheme adjusts their magnitude to meet accuracy requirements. To handle high-dimensional settings efficiently, we replace full eigenvalue decompositions with a Rayleigh quotient approximation, focusing only on the critical invariant subspace needed to restore positive semidefiniteness.

The resulting algorithm systematically incorporates structural constraints into the nearest correlation matrix problem. Numerical experiments highlight its robustness across diverse structured scenarios, with promising applications in finance, statistics, and network analysis.

Session 2 / **7**

Emergence and Stability of Deep Neural Collapse

Authors: Emanuele Zangrando[1]; Francesco Tudisco[2]; Nicola Guglielmi[1]; Piero Deidda[1]; Simone Brugiapaglia[3]

[1] *Gran Sasso Science Institute*

[2] *GSSI*

[3] *Concordia University*

Corresponding Author: emanuele.zangrando@gssi.it

Recent empirical and theoretical results suggest that deep networks possess an implicit low-rank bias: their weight matrices naturally evolve toward approximately low-rank structure, and a structured pruning of small singular values can often reduce model size with little or no loss in accuracy. While this phenomenon is already understood in simplified settings, a complete theory accounting for the effects of nonlinearities is still missing.
In this talk, we will present a framework that connects deep neural collapse to the emergence of low-rank structure in a broad class of nonlinear feedforward networks. For both nonlinear feedforward and residual architectures, we prove the global optimality of collapsed solutions and show that interpolating minima are effectively barrier-free paths to these global optima, offering a possible explanation for the ubiquity of collapse in practice.

Session 5 / **8**

Past-aware game-theoretic centrality in complex contagion

Author: Francesco Zigliottto[1]

[1] *Scuola Normale Superiore*

Corresponding Author: francesco.zigliotto@sns.it

We introduce past-aware game-theoretic centrality, a class of centrality measures that captures the collaborative contribution of nodes in a network, accounting for both uncertain and certain collaborators. A general framework for computing standard game-theoretic centrality is extended to the past-aware case. As an application, we develop a new heuristic for different versions of the influence maximization problem in complex contagion, which models processes requiring reinforcement from multiple neighbors to spread. A computationally efficient explicit formula for the corresponding past-aware centrality score is derived, leading to scalable algorithms for identifying the most influential nodes, which in most cases outperform the standard greedy approach in both efficiency and solution quality.

Session 2 / **9**

Accuracy-cost trade-offs & Efficiency in Neural Physics Simulation

Author: Pietro Sittoni[1]

Co-authors: Francesco Tudisco [2]; Nicola Guglielmi [3]; Emanuele Zangrando [3]; Angelo Alberto Casulli [4]; Harris Abdul Majid [5]

[1] *Gran Sasso Science Institute- GSSI*

[2] *GSSI*

[3] *Gran Sasso Science Institute*

[4] *Scuola Normale Superiore di Pisa*

[5] *The University of Edinburgh*

Corresponding Author: pietro.sittoni@gssi.it

Accuracy–cost trade-offs and data and parameter efficiency are two fundamental aspects of machine learning for scientific computing. In the first part of this talk, we address test-time control of model performance. We introduce the Recurrent-Depth Simulator (RecurrSim), an architecture-agnostic framework that enables explicit control over accuracy–cost trade-offs in neural simulators without retraining or architectural redesign. By adjusting the number of recurrent iterations, users can flexibly trade computational cost for accuracy at inference time. RecurrSim achieves physically plausible long-horizon simulations across standard fluid dynamics benchmarks and large-scale 3D compressible Navier–Stokes problems, where a 0.8B-parameter RecurrFNO outperforms 1.6B-parameter baselines while using 13.5% less training memory. The framework generalizes across diverse architectures, including transformers and operator-learning models.
In the second part of the talk, we focus on data and parameter efficiency. We introduce Neural-HSS, a novel architecture, inspired by the structure of Green's functions for elliptic PDEs, based on Hierarchical Semi-Separable matrices. Neural-HSS is provably data-efficient, satisfies exactness properties in low-data regimes for a broad class of PDEs, and empirically demonstrates superior performance on large-scale elliptic and multi-physics PDEs across diverse scientific domains.

Session 3 / **10**

Sparse identification of delay equations with distributed memory

Author: Muhammad Tanveer[1]

Co-authors: Dimitri Breda [1]; Jianhong Wu [2]

[1] *University of Udine, Italy*

[2] *York University, Toronto, Canada*

Corresponding Author: tanveer.muhammad@spes.uniud.it

Data-driven model discovery has become a powerful approach for identifying governing equations of dynamical systems using temporal data. The Sparse Identification of Nonlinear Dynamics (SINDy) algorithm, initially developed for ordinary differential equations (ODEs), has been extended to more general classes of problems, including partial differential equations (PDEs), stochastic differential equations (SDEs) and delay differential equations (DDEs) with discrete delays. However, its application to systems with distributed delays and renewal equations remains unexplored.
Distributed delays involve integration over a continuum of past states and are at the core of renewal-type integral equations. Related models are prevalent in biological and ecological applications, to describe in particular structured populations and epidemics. They portrait memory dependent dynamics but are challenging to identify due to the inherent complexity of delay kernels and renewal processes. Building on the integral formulation of SINDy for ODEs, we propose a novel extension of

the SINDy framework to recover the (possibly nonautonomous) kernel of distributed delays through the use of quadrature formulas. As such the new approach aims at providing a sparse interpretable model rather than just a black-box right-hand side.
We demonstrate the efficacy of the new method first on academic examples and then by applying it to the transmission dynamics of Severe Fever with Thrombocytopenia Syndrome (SFTS), an emerging tick-borne disease. We use real data on reported human cases and daily temperature from Dalian, China. The technique successfully discovers a parsimonious integral model that not only accurately predicts the outbreaks but also identifies an interpretable, temperature dependent kernel. This study validates our extended SINDy framework as a powerful tool for uncovering and forecasting the dynamics of complex environmental and epidemiological systems governed by memory effects.

Session 4 / **11**

One probe is enough for log-det estimation

Authors: Alice Cortinovis[1]; Daniele Toni[2]

[1] *Università di Pisa*

[2] *Scuola Normale Superiore*

Corresponding Author: alice.cortinovis@unipi.it

The log-determinant of a symmetric positive semi-definite matrix is a quantity that arises in different contexts, for instance in the evaluation of the log-marginal likelihood for Gaussian processes and in the normalization of the determinantal point processes for supervised learning.
We focus on randomized algorithms for estimating this quantity. The algorithms access the matrix only through matrix vector products, and are based on the introduction of a preconditioner and stochastic trace estimator.
We claim that preconditioning as much as we can and making a rough estimate of the residual part with a small budget achieves a small error in most of the cases. We choose a Nystr\"om preconditioner and estimate the residual using only one sample of stochastic Lanczos quadrature. We analyze the performance of this strategy from a theoretical and practical viewpoint. Numerical examples on several test matrices show that our proposed method is competitive with existing algorithms.

Session 5 / **13**

Area-Preserving Surface Parameterizations via Riemannian Optimization

Author: Marco Sutti[1]

Co-author: Mei-Heng Yueh [2]

[1] *GSSI*

[2] *National Taiwan Normal University*

Corresponding Author: marco.sutti@gssi.it

I will present two recent works, coauthored with Mei-Heng Yueh, in which we propose Riemannian optimization algorithms for computing spherical and toroidal area-preserving mappings of genus-zero and genus-one closed surfaces, respectively. The proposed framework is based on retraction-based Riemannian optimization, which provides an effective way to handle the geometric constraints of the problem.
I will first introduce the main components of the Riemannian optimization framework and discuss

the objective function and underlying geometry. I will then present numerical experiments on several mesh models, demonstrating the effectiveness of the proposed approach compared with existing state-of-the-art methods for computing area-preserving mappings. Finally, I will illustrate applications of the proposed algorithms to landmark-aligned surface registration and texture mapping.

Session 5 / **14**

Long-time relative error analysis for linear ordinary differential equations with perturbed initial value

Author: Stefano Maset[1]

[1] *Università degli Studi di Trieste*

Corresponding Author: maset@units.it

We investigate the propagation of initial value perturbations along the solution of a linear ordinary differential equation $y'(t) = Ay(t)$. This propagation is analyzed using the relative error rather than the absolute error. Our focus is on the long-term behavior of this relative error, which differs significantly from that of the absolute error. Understanding the long-term behavior provides insights into the growth of the relative error over all times, not just at large times.

Session 4 / **15**

New Algorithms to Solve Non-Linear Least Square Problems

Authors: Francesco Tudisco[None]; Nicola Guglielmi[None]

Corresponding Author: asma.farooq@gssi.it

Non-linearity often leads to slow or unstable convergence in iterative solvers for nonlinear least-squares problems. In this work, we introduce a family of accelerated algorithms that leverage a periodically restarted variant of the Generalized Minimum Residual (GMRES) method to address these challenges. The restarting strategy keeps the computational cost under control and makes the method more reliable when dealing with nonlinear problems.
To further improve efficiency, our approach integrates sketching and subspace-projection mechanisms, enabling substantial reductions in computational cost while maintaining high solution accuracy. We derive an explicit perturbation bound that quantifies how inaccuracies in the reduced subspace influence the stability of the computed solution.

Extensive numerical experiments confirm that the resulting algorithms deliver accurate reconstructions with markedly improved computational performance compared to classical gradient-based techniques.

Session 1 / **16**

On superdiffusive processes on graphs using a regularized fractional Laplacian

Author: Alessandro Filippo[1]

Co-author: Mariarosa Mazza [2]

[1] *Università degli studi di Roma Tor Vergata*

[2] *University of Rome Tor Vergata*

Corresponding Author: alefil1@outlook.it

The fractional graph Laplacian, defined as a fractional power of the standard graph Laplacian, is one of the most popular tools for modeling non-local diffusion on graphs. However, it is known to induce dynamics that, in some cases, are incompatible with the topology of the original network. To address this limitation, a regularized fractional Laplacian obtained through a combination of the standard and fractional Laplacians, was recently introduced with the aim of restoring compatibility while preserving the spectral richness of the fractional operator.

In this talk, we explore the spectral and diffusion properties of the regularized fractional Laplacian. Contrary to the fractional Laplacian and other non-local Laplacian variants, we show that it consistently produces superdiffusive behavior, regardless of whether the underlying graph is weighted or unweighted. We also present an efficient Boolean–Hadamard–based construction of the regularized operator that improves its computational practicality in real-life scenarios.

Session 1 / **18**

Discrete-To-Continuum Limits in Graph-Based Semi-Supervised Learning

Author: Matthew Thorpe[1]

[1] *University of Warwick*

Corresponding Author: matthew.thorpe@warwick.ac.uk

Semi-supervised learning (SSL) is the problem of finding missing labels from a partially labelled data set. The heuristic one uses is that "similar feature vectors should have similar labels". The notion of similarity between feature vectors explored in this talk comes from a graph-based geometry where an edge is placed between feature vectors that are closer than some connectivity radius. A natural variational solution to the SSL is to minimise a Dirichlet energy built from the graph topology. And a natural question is to ask what happens as the number of feature vectors goes to infinity? In this talk I will give results on the asymptotics of graph-based SSL using an optimal transport topology. The results will include a lower bound on the number of labels needed for consistency and, time permitting, some recent extensions to infinite dimensional settings.

Session 3 / **19**

Learning mechanical systems from real world data

Author: Elena Celledoni[1]

[1] *NTNU*

Corresponding Author: elena.celledoni@ntnu.no

In this talk I will consider a method for learning the Lagrangian and forces for mechanical systems using the discrete Lagrange d'Alembert principle. The case of manifold valued data and data on Lie groups will also be discussed if time permits.

I will also describe a number of current projects with diverse applications where the main theme is learning vector fields from data.

Session 1 / **20**

Fractional diffusion on a graph as a memory-driven diffusion

Authors: Nikita Deniskin[1]; Ernesto Estrada[2]

[1] *Scuola Normale Superiore*

[2] *IFISC - UIB*

Corresponding Author: nikita.deniskin@sns.it

Subdiffusion on networks can occur where overcrowding is present. It is studied through time-fractional diffusion equations, and admits an explicit solution through the Mittag-Leffler function. We give a representation of subdiffusion as a superposition of classical diffusion processes with subordination to a different timescale. Memory arises in subdiffusion, while the classical diffusion processes are Markovian.

We study the subdiffusive geometry of the network that is uncovered by the dynamics, a new metric on the graph obtained with the Communicability Distance. Fixing two vertices, the shortest paths between them are vastly different in subdiffusion and diffusion geometries. We show how memory emerges from these two different behaviours and the relation with operator-valued Volterra evolution equations.

23

Opening Remarks

Opening Remarks / Welcome & Opening

25

Concluding Remarks

Session 2 / **27**

Signal processing of spherical data. How manifold topology interacts with data analysis techniques

Author: Roberto Cavassi[1]

Co-authors: Edward J. Timko ; Giovanni Barbarino ; Haomin Zhou ; Wing S. Li

[1] *University of L'Aquila*

In signal processing, the time-frequency analysis of nonlinear and non-stationary processes, as well as the determination of the unknown number of active sub-signals in a blind-source composite signal, are generally challenging inverse problem tasks. If we consider data sampled on a sphere, things get even more complicated. This is the reason why just a few techniques have been developed so far to study this kind of data. However, many real-life data are of this nature, like in Geophysics and Physics.
The idea is to extend the Iterative Filtering (IF) algorithm to work on the sphere. IF is a non-stationary

signal decomposition method proposed a decade ago [Lin et al. 2009] to address the problem of extracting time-varying oscillatory components from a non-stationary multicomponent signal. This method proved to be really valuable in many applications, see [Barbarino&Cicone 2022] and references therein, and it was accelerated in what is known as Fast Iterative Filtering (FIF) [Cicone&Zhou 2021] by leveraging the Toeplitz matrix theory. In this talk, we introduce the generalization of IF to handle spherical data and show how we can address the question about its convergence [Barbarino et al 2024]. We conclude with some examples of application to geophysical data.

L. Lin, Y. Wang, and H. Zhou. Iterative filtering as an alternative algorithm for empirical mode decomposition. Adv. in Adap. Data An., 2009, 1.04, 543-560.

G. Barbarino, A. Cicone. Conjectures on spectral properties of ALIF algorithm. Linear Algebra and its Applications, Volume 647, Pages 127-152, 2022.

A. Cicone, H. Zhou. Numerical Analysis for Iterative Filtering with New Efficient Implementations Based on FFT. Num. Math., 2021, 147 (1), 1–28.

G. Barbarino, R. Cavassi, A. Cicone. Extension and convergence analysis of Iterative Filtering to spherical data. Linear Algebra and its Applications, 2024.

Session 3 / **28**

Manifold-based Algorithms for the Hadamard Decomposition

Author: Stefano Sicilia[1]

[1] *University of Mons*

Corresponding Author: stefano.sicilia@umons.ac.be

The Hadamard decomposition is a powerful technique for data analysis and compression. Given an $m \times n$ matrix X and two ranks $r_1, r_2 \ll \min(m, n)$, we look for two low-rank matrices X_1 and X_2 of the same size of X and with $\text{rank}(X_i) = r_i$ such that $X \approx X_1 \circ X_2$, where $\circ$ is the element-wise product. In contrast with the well-known SVD, this decomposition allows to represent higher-rank matrices with the same amount of variables. Remarkably, it can be proved that any Hadamard factorization for X can be rewritten as $X \approx WH^\top$ with W of rank r_1 and and H of rank-r_2. In this work we focus on the most interesting case where $r = r_1 = r_2$; we present some new theoretical results which show the features of this factorization and when it is efficient. Based on these facts, we develop two new manifold-based gradient descent algorithms for computing the rank-r Hadamard decomposition: the first one concerns the representation $X \approx X_1 \circ X_2$, while the second one focuses on $X \approx WH^\top$ and it integrates a gradient system associated to the factors W and H. We also derive some new initialization guesses that could be used for any iterative method that solves the problem. We compare our results with the SVD and with one of the few existing approaches in the literature [Wertz et al., 2025] to compute the rank-r Hadamard decomposition. Numerical results prove that the new methods are efficient and competitive, both on synthetic and real data. We also show that the new initializations proposed usually improve the performances of the algorithms.

Session 2 / **29**

Inverse problems in signal processing: Old problems and new opportunities

Author: Antonio Cicone[1]

[1] *University of L'Aquila*

Corresponding Author: antonio.cicone@univaq.it

In many applied research fields—including Geophysics, Medicine, Engineering, Economics, and Finance— fundamental challenges involve extracting hidden information and meaningful features from complex signals, such as quasi-periodicities, time-varying frequency patterns, and underlying components like trends.

Classical signal processing techniques based on Fourier and Wavelet transforms, while powerful, often exhibit intrinsic limitations when dealing with nonlinear and strongly non-stationary phenomena. Over the past two decades, this has motivated the development of several adaptive, data-driven nonlinear methods, which have been successfully applied across a wide range of scientific disciplines.

In this talk, we first briefly review the Hilbert–Huang Transform (also known as Empirical Mode Decomposition) and discuss its main theoretical and practical limitations. We then introduce the Iterative Filtering framework and its extensions for the analysis of multidimensional, multivariate, and highly non-stationary signals, together with the recently proposed time–frequency representations known as the IMFogram and JADE algorithms. We will illustrate their theoretical properties and numerical performance through applications to real-world data. We will also present open challenges that await being tackled in this field of research.

Session 4 / **30**

Efficient Krylov-like solvers for model and data driven problems

Author: Silvia Gazzola[None]

Corresponding Author: silvia.gazzola@unipi.it

This talk is about efficient solution methods for inverse problems, i.e., the task of recovering an object of interest from known but corrupted data available through a known but possibly corrupted model, formulated in a discrete and linear setting. Since these problems are ill posed, variational regularization methods are often applied to recover a meaningful solution. However, variational regularization methods always require a number hyperparameters, i.e., parameters that must be pre-specified in advance, such as regularization parameters. A data-driven approach to determine appropriate hyperparameter values is via a nested optimization framework known as bilevel learning. This talk will focus on novel strategies, based on Krylov methods, to effectively and efficiently solve such bilevel learning problems. Namely: (1) The use of non-standard Krylov methods to solve the regularized problem will be presented, to set a variety of hyperparameters (in addition to the regularization parameters). (2) The application of recycling Krylov subspace methods will be explored, resulting in cheaper computations of each hypergradient in a gradient-based optimization method for solving the bilevel learning problem. The proposed approaches are validated through extensive numerical testing in the context of inverse problems in imaging.