Ph.D. in Information Technology
Thesis Defenses

May 11th, 2026
At 9:30 a.m.
Sala Schiavoni - Building 20A

Federico MIOTELLO – XXXVIII Cycle

MODELS AND METHODS FOR SOUND FIELD ANALYSIS AND RECONSTRUCTION

Supervisor: Prof. Alberto Bernardini

Abstract:

The ability to capture, manipulate, and reproduce sound fields with spatial fidelity is a fundamental requirement of modern acoustic signal processing. Applications such as immersive multimedia, virtual and augmented reality, teleconferencing, spatial audio rendering, and intelligent acoustic sensing all rely on accurate models of how sound propagates in space and how it interacts with the devices that capture or radiate it. As these devices become increasingly compact and computational resources more powerful, there is a growing need for methodologies that combine physically grounded acoustic modeling with advanced signal processing and learning-based inference techniques.

Within this broad scenario, this thesis investigates models and methods for sound field analysis and reconstruction under practical hardware and sampling constraints. The work addresses two complementary problems that are central to modern spatial audio systems: robust spatial filtering with compact arrays and the reconstruction of spatially extended acoustic fields from sparse measurements. Although they pursue different goals and have traditionally evolved as distinct research lines, both of them can be formulated as structured inverse problems whose solutions depend on assumptions that classical formulations tend to idealize: the directivity of array elements, the density of spatial sampling, the priors available for inference. Revisiting these assumptions, and replacing them with more faithful models of transducer behavior, explicit physical constraints, or learned spatial priors, is what enables practical systems to perform well under realistic conditions.

The first part of the thesis focuses on differential beamforming with compact microphone and loudspeaker arrays. Differential arrays are particularly attractive because they enable highly directional and nearly frequency-invariant spatial responses while operating with apertures significantly smaller than the wavelength of interest. This property makes them well suited for broadband audio applications in embedded, portable, and miniaturized systems. However, classical differential beamforming theory relies on the simplifying assumption that array elements behave as ideal omnidirectional transducers. In practice, small-size microphones and loudspeakers exhibit intrinsic, often frequency-dependent, directivity patterns that may substantially deviate from this assumption. First, the thesis provides a systematic analysis of the impact of array element directivity on differential beamformer performance. Building upon this analysis, steerable first-order differential loudspeaker arrays composed of monopole and dipole elements are introduced, demonstrating how

element directionality can be explicitly exploited rather than neglected. Finally, a generalized modal-matching framework is proposed for frequency-invariant differential beamforming with arbitrary planar arrays of first-order directional elements.

The second part of the thesis addresses sound field reconstruction under sparse spatial sampling. Two complementary strategies are investigated. The first formulates reconstruction as a conditional generative problem using diffusion models, enabling high-quality estimation of acoustic field magnitudes from limited observations by leveraging learned spatial priors. The second introduces a zero-shot, physics-informed dictionary learning method that enforces the Helmholtz equation during optimization, yielding physically consistent reconstructions without requiring large training datasets. In addition, a physics-informed neural network approach for spatial upsampling of spherical microphone arrays is presented, allowing compact arrays to emulate higher-order systems while preserving physical plausibility.

Overall, the thesis demonstrates that accurate modeling of the directivity of acoustic transducers, explicit enforcement of wave-equation constraints, and principled integration of data-driven priors are key to advancing space-time acoustic processing. The contribution of this work consists of a set of methodologies that advance the state of the art in differential beamforming and sound field reconstruction, each grounded in physical modeling and tailored to realistic hardware and sampling conditions.

Oliviero MASSI– XXXVII Cycle

DIFFERENTIABLE CIRCUITAL MODELS OF AUDIO SYSTEMS FOR DIGITAL SIGNAL PROCESSING

Supervisor: Prof. Alberto Bernardini

Abstract:

Audio systems, ranging from analog effects and guitar amplifiers to modern MEMS loudspeakers, are inherently multiphysics and nonlinear devices. The nonlinear interactions among different physical domains, together with phenomena such as magnetic or piezoelectric hysteresis, are key to the distinctive behavior and sonic character of these systems. Capturing these effects in a lightweight digital representation is essential for both Virtual Analog (VA) modeling, which aims to faithfully emulate analog audio circuits in the digital domain, and for the design and optimization of audio transducers, where accurate physical models support the development of Digital Signal Processing (DSP) algorithms for performance enhancement. Lumped-Element Models (LEMs) provide a compact and physically interpretable framework for describing such systems by exploiting analogies between physical domains and representing them as equivalent electrical networks of interconnected circuit elements. These models are particularly well suited to audio applications, as they can be implemented efficiently in the discrete-time domain and interfaced with DSP structures. However, LEMs have limited expressive power when it comes to capturing strongly nonlinear or hysteretic effects, and their analytical treatment often leads to implicit equations that are difficult to solve in real-time. To address these challenges, this dissertation extends the Wave Digital Filter (WDF) formalism into a unified, differentiable modeling and implementation framework. In this approach, circuit elements and connection networks are described as input-output blocks characterized by scattering relations, where port voltages and currents are expressed as linear combinations of incident and reflected waves through the introduction of a reference port resistance. The choice of this free parameter per port provides important numerical advantages, ensuring stability, modularity, and computational efficiency. Building on this modular structure, in this thesis we introduce neural network-based Wave Digital (WD) blocks for representing the nonlinear scattering equations of one-port and multi-port elements. Both explicit and implicit hybrid WD structures are investigated: explicit realizations enable real-time simulation of circuits containing a single nonlinear neural element, while implicit realizations address more complex circuits with multiple nonlinear WD blocks using iterative solvers. Both standard and hybrid WD structures can be simulated in the discrete-time domain within an Automatic Differentiation (AD)-based framework, which enables the computation of exact derivatives of the simulated response with respect to the model parameters. This allows the gradient-based optimization of both circuit parameters and neural parametric models directly from experimental measurements. In this setting, the parameters of linear and nonlinear LEMs can be refined according to a supervised learning paradigm, with neural WD blocks trained within the same differentiable simulation loop. The proposed modeling framework is applied to two main areas. In VA modeling, hybrid WDF structures are used to emulate nonlinear audio circuits achieving high accuracy and real-time efficiency; in piezoelectric MEMS loudspeaker modeling, differentiable LEMs combining WDF implementations and neural network-based processing blocks are used to capture the nonlinear electro-mechano-acoustic behavior of these transducers, supported by data-driven parameter identification directly from acoustic measurements. Finally, building upon the established LEM modeling framework, we address circuit inversion and transducer virtualization, introducing a systematic method for deriving inverse circuit representations based on nullors. This enables the design of model-based DSP algorithms that digitally compensate or modify the behavior of audio transducers through cascaded direct and inverse systems. The approach is applied to a piezoelectric MEMS loudspeaker equalization case study, demonstrating accurate response flattening while preserving the interpretability of the underlying physical model.

PhD Committee

Dr. Marco Marcon, **Politecnico di Milano**

Prof. Federico Fontana, **Università degli Studi di Udine**

Prof. Antonio Jesus Muñoz Montoro, **Universidad de Jaen**