

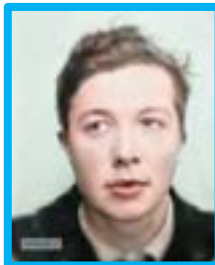
Lie Groups Statistics and Machine Learning for Military Sensors based on Symplectic Structures of Information Geometry

Frédéric BARBARESCO

THALES KTD PCC SENSING SEGMENT LEADER

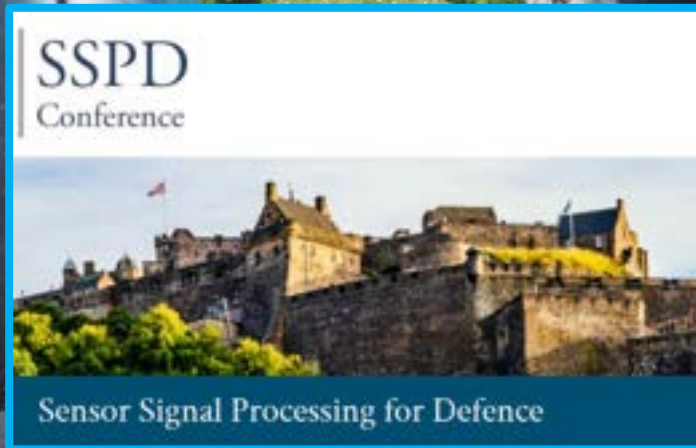
(Key Technology Domain « Processing, Control & Cognition »)

Jean-Marie
SOURIAU



Jean-Louis
KOSZUL

OPEN



Advanced Sensor Processing based on Koszul Information Geometry

- Geometric Matrix CFAR/STAP for (very) slow targets detection in clutter
- Complex-Valued CNN & Covariance-Matrix-Valued HPDNet for Micro-Doppler ATDR
- Tracker parameters tuning by Deep Learning for tracking hyper-maneuvering targets
- Multi-Agent Reinforcement Learning for Sensor Resources Management for tracking hyper-maneuvering targets
- Multi-Sensors Collaborative Tracking by Distributed Auctions for tracking in saturating scenario (swarm, fleet of targets, ...)

Modern Sensor Processing based on Symplectic Model of Information

- Lie Group Based Equivariant GCNN for Adaptive Doppler Clutter Map
- Lie Group Based Frenet-Serret IEKF (Invariant Extended Kalman Filter) for tracking hyper-maneuvering targets
- Lie Group Based Target Recognition on Kinematics for Drone/Birds Classification
- Souriau Symplectic Model of Information for Lie Group Statistics and Machine Learning
 - Entropy as Casimir Function in Coadjoint Representation
 - Koszul-Fisher Metric on Lie Group
 - Covariant Maximum Entropy Density (Gauss Density) on Lie Group
 - Lie Groups Machine Learning

From PASCALINE Machine to HPC or Geometric Integrating Machines



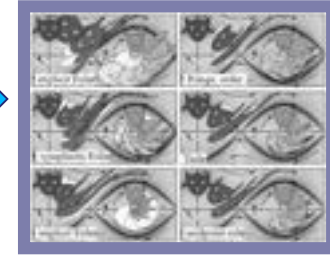
Pascaline

Analytic Machines
based on Linear Algebra



Jacquard Loom

Geometric
Machines



Geometric & Symplectic
Integrators based on Lie
Group Algebra
(Intrinsic Computation without
coordinates)



Babbage Analytic
Machine



Descartes
computation with
coordinates



High Power
Computing

HALES
a future we can all trust

OPEN

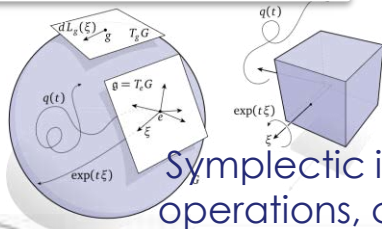
AI/Machine Learning Evolution: ALGEBRA COMPUTATION STRUCTURES

Calcul formel pour les méthodes de Lie en mécanique hamiltonienne

P.V. Koseleff, X/CMLS PhD, 1993 (P. Cartier)

Souriau Exponential Map Algorithm for Machine Learning on Matrix Lie Groups

Frédéric Barbaresco, Springer GSI'19, 2019



Symplectic integrators, non-commutative operations, coadjoint orbits, moment map

Supervarieties, Sow. Math. Dokl. 16 (1975), 1218-1222.

F. A. Berzin and D. A. Leites

$$X = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

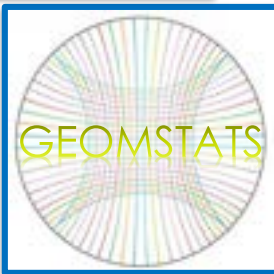
$\text{Ber}(X) = \det(A) \det(D - CA^{-1}B)^{-1}$
Berezian Determinant

LIE ALGEBRA

LIE SUPER ALGEBRA



Computer Algebra Group- Scratchpad, IBM, 1971



$$\begin{bmatrix} A & B \\ C & D \\ E & F \end{bmatrix} \times \begin{bmatrix} G \\ H \end{bmatrix} = \begin{bmatrix} A \times G + B \times H \\ C \times G + D \times H \\ E \times G + F \times H \end{bmatrix}$$

LINEAR ALGEBRA

Vectors space, commutative matrix operations, eigen-analysis

BOOLE ALGEBRA

Boolean logic digital circuits using electromechanical relays as the switching element.



GOOGLE TPU (Tensor Processing Unit)

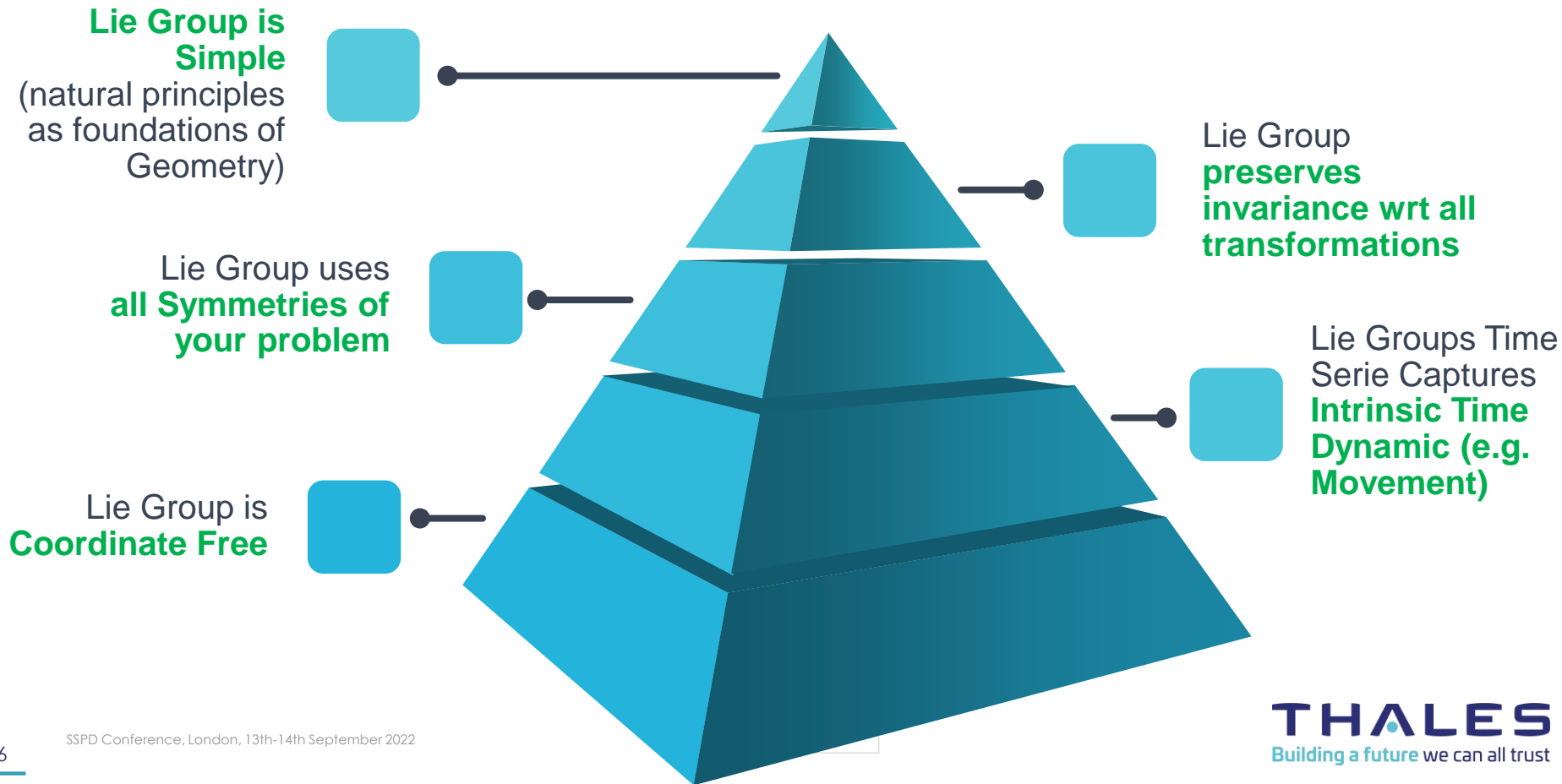


George R. Stibitz (Bell Labs)

ALGEBRA is the study of mathematical symbols and the rules for manipulating these symbols

Building a future we can all trust

Rational to Use Lie Groups for THALES Machine Learning Applications

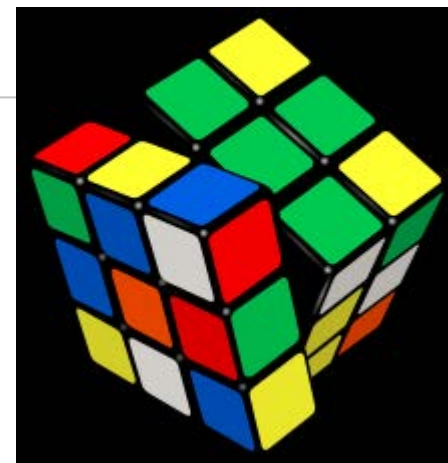


Lie Group

GROUP (Mathematics)

A set equipped with a binary operation with 4 axioms:

- Closure $\forall a, b \in G$ then $a \bullet b \in G$
- Associativity $\forall a, b, c \in G$ then $(a \bullet b) \bullet c = a \bullet (b \bullet c)$
- Identity $\exists e \in G$ such that $e \bullet a = a \bullet e = a$
- invertibility $\forall a \in G, \exists b \in G$ such that $b \bullet a = a \bullet b = e$



LIE GROUP

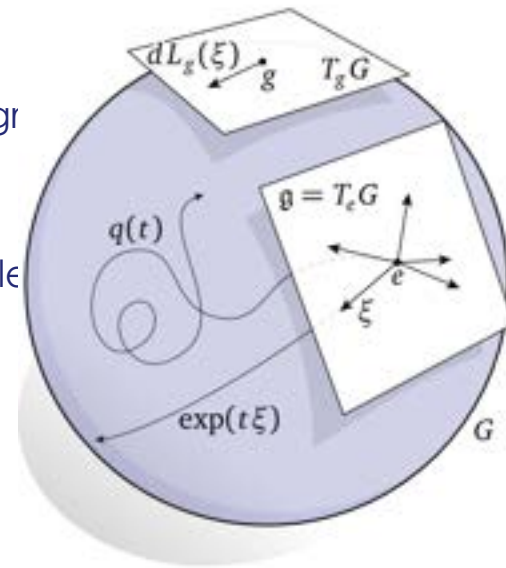
- A group that is a differentiable manifold, with the property that the group operations of multiplication and inversion are smooth maps:

$\forall x, y \in G$ then $\phi: G \times G \rightarrow G$ then $\phi(x, y) = x^{-1}y$ is smooth

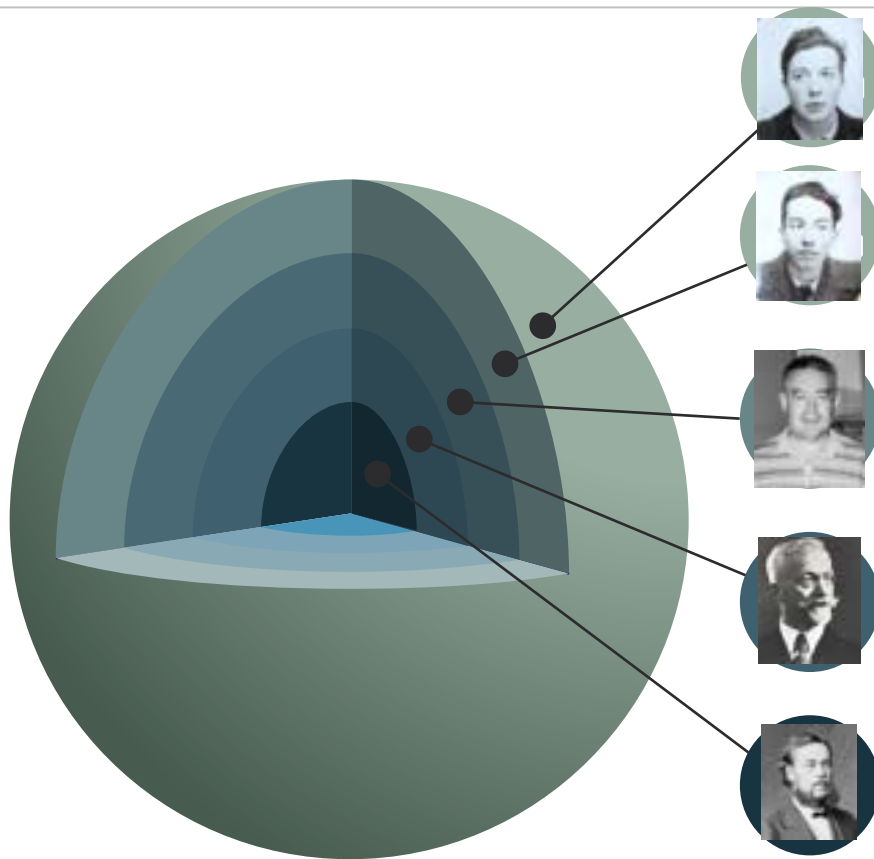
- A Lie algebra $\mathfrak{g} = T_e G$ is a vector space with a binary operation called the Lie bracket $[\cdot, \cdot]: \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ that satisfies axioms:

$$[ax + by, z] = a[x, z] + b[y, z] ; [x, x] = 0 ; [x, y] = -[y, x]$$

$$\text{Jacobi Identity: } [x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0$$



Lie Groups Tools Development: From Group to Co-adjoint Orbits



Lie Group & Statistical Physics

Jean-Michel Bismut – Random Mechanics

Jean-Marie Souriau – Lie Group Thermodynamics, Souriau Metric

Jean-Louis Koszul – Affine Lie Group & Algebra representation

Harmonic Analysis on Lie Group & Orbits Method

Pierre Torasso & Michèle Vergne – Poisson-Plancherel Formula

Michel Duflo – Extension of Orbits Method, Plancherel & Character

Alexandre Kirillov – Coadjoint Orbits, Kirillov Character

Jacques Dixmier – Unitary representation of nilpotent Group

Lie Group Representation

Bertram Kostant – KKS 2-form, Geometric Quantization

Alexandre Kirillov – Representation Theory, KKS 2-form

Jean-Marie Souriau – Moment Map, KKS 2-form, Souriau Cocycle

Valentine Bargmann – Unitary representation, Central extension

Lie Group Classification

Carl-Ludwig Siegel – Symplectic Group

Hermann Weyl – Conformal Geometry, Symplectic Group

Elie Cartan – Lie algebra classification, Symmetric Spaces

Willem Killing – Cartan-Killing form, Killing Vectors

Group/Lie Group Foundation

Henri Poincaré – Fuchsian Groups

Felix Klein – Erlangen Program (Homogeneous Manifold)

Sophus Lie – Lie Group

Evariste Galois/Louis Joseph Lagange – Substitution Groups

OPEN

Structuring Principles for Learning : Calculus of Variations

d. Translated, in any way, in whole or in part
- © 2021 THALES. All rights reserved.



Pierre
de Fermat

Fermat's principle
of least time

Pierre
Louis
Maupertuis



Maupertuis's
principle of
least length

Joseph
Louis
Lagrange



(Euler)
Lagrange
Equation

Simeon
Denis
Poisson



Poisson
Bracket,
Poisson Geometry
Structure

Henri
Poincaré



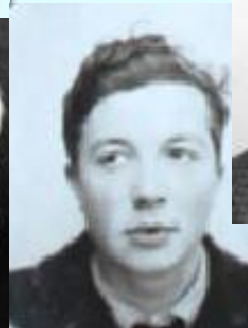
(Euler)
Poincaré
Equation

Elie
Cartan



Poincaré
Cartan
Integral
Invariant

Jean- Marie
Souriau



Souriau
Moment
Map,
Souriau
Symplectic
2 Form,
Lie Groups
Thermodynamics

Jean-Michel
Bismut



Random
Mechanics

Bedrock of Symplectic Theory of Information Geometry

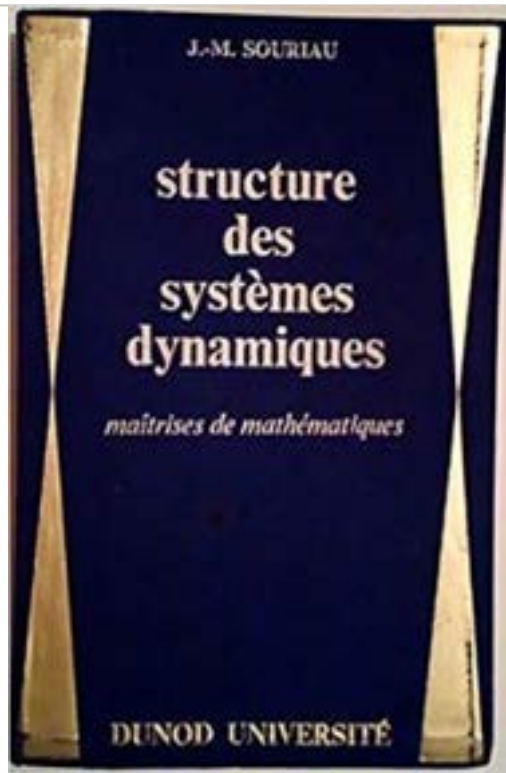
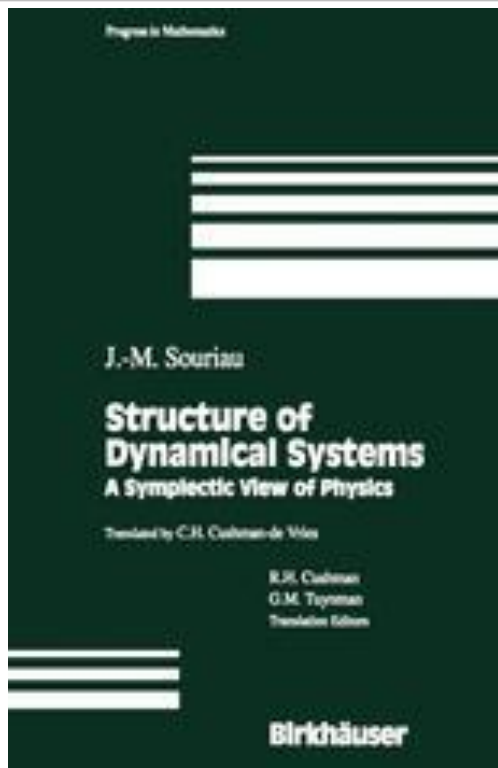


Jean-Marie Souriau (ENS 1942)



Jean-Louis Koszul (ENS 1940)

J.M. Souriau Book « Structure des systèmes dynamiques », 1969



- Introduction of symplectic geometry in mechanics
- Invention of the “moment map”
- Geometrization of Noether's theorem
- Barycentric decomposition theorem
- The total mass of an isolated dynamic system is the class of cohomology of the default of equivariance for the moment map
- Lie Groups Thermodynamics (Chapter IV)

http://www.jmsouriau.com/structure_des_systemes_dynamiques.htm
<http://www.springer.com/us/book/9780817636951>

Internet website : <http://souriau2019.fr>

In 1969, 50 years ago, Jean-Marie Souriau published the book "**Structure des système dynamiques**", in which using the ideas of J.L. Lagrange, he formalized the "**Geometric Mechanics**" in its modern form based on **Symplectic Geometry**

Chapter IV was dedicated to "Thermodynamics of Lie groups" (ref André Blanc-Lapierre)

Testimony of **Jean-Pierre Bourguignon** at Souriau'19 (IHES director of the European ERC)



*Jean-Marie SOURIAU
and
Symplectic Geometry*

*Jean-Pierre BOURGUIGNON
(CNRS-IHÉS)*

<https://www.youtube.com/watch?v=93hFollBo0Q&t=3s>



JEAN-MARIE SOURIAU

In 1969, the groundbreaking book of Jean-Marie Souriau appeared "Structure des Systèmes Dynamiques". We will celebrate, in 2019, the jubilee of its publication, with a conference in honour of the work of this great scientist.

Symplectic Mechanics, Geometric Quantization, Relativity, Thermodynamics, Cosmology, Diffeology & Philosophy

Frédéric Barbaresco
 David Barina
 Jean-Pierre Bourguignon
 Pierre Cartier
 Gipi Christodoulou
 Maurice Coste
 Thibault Dambrin
 Paul Delorme
 Paolo Donati
 Daniel Goebl
 Patrick Iglesias-Arnáez
 Jean-Pierre Magnin
 Pierre-Michel Schneider
 Marc Lachapelle-Ray
 Marie Perle
 Elise Poincaré
 Jean-Jacques Stieglitz
 Richard Tey
 André Weil
 Jean-Michel Tziou
 Alan Weinstein

80|Prime



HOMOGENEOUS SYMPLECTIC SPACES AND CENTRAL EXTENSIONS

MAXENT2022

ANDREW BECKETT
(WORK WITH JOSÉ FIGUEROA-O'FARRILL)

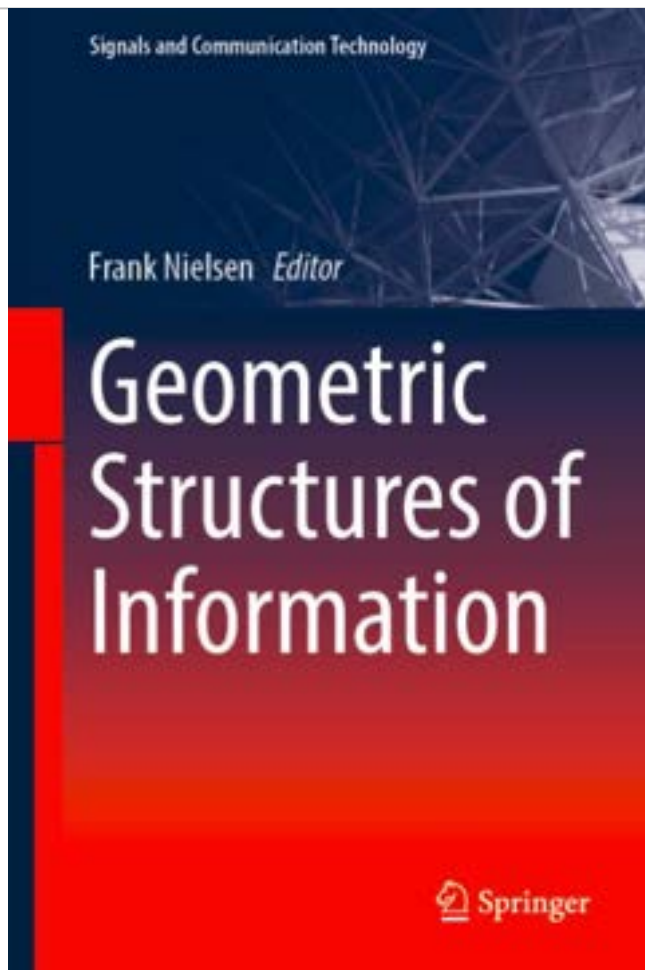
MAXWELL INSTITUTE GRADUATE SCHOOL,
UNIVERSITY OF EDINBURGH

21 JULY 2022



MAXWELL INSTITUTE FOR
MATHEMATICAL SCIENCES





Geometric Structures of Information

➤ <https://www.springer.com/us/book/9783030025199>

Paper on Jean-Louis Koszul

➤ Barbaresco, F. , Jean-Louis Koszul and the Elementary Structures of Information Geometry, Geometric Structures of Information, pp 333-392, SPRINGER, 2018

➤ https://link.springer.com/chapter/10.1007%2F978-3-030-02520-5_12

Koszul Book on Souriau Work: The Little Green Book



Koszul Book on Souriau Work: The Little Green Book

Jean-Louis Koszul · Yiming Zou

Introduction to Symplectic Geometry

Forewords by Michel Ngulffo Boyom, Frédéric Barbaresco and Charles-Michel Marle

This introductory book offers a unique and unified overview of symplectic geometry, highlighting the differential properties of symplectic manifolds. It consists of six chapters: Some Algebra Basics, Symplectic Manifolds, Cotangent Bundles, Symplectic G-spaces, Poisson Manifolds, and A Graded Case, concluding with a discussion of the differential properties of graded symplectic manifolds of dimensions (o, n) . It is a useful reference resource for students and researchers interested in geometry, group theory, analysis and differential equations.

$$\mu : M \longrightarrow \mathfrak{g}^*$$

$$\mu(sx) = s\mu(x) = \text{Ad}^*(s)\mu(x) + \varphi_\mu(s), \quad \forall s \in G, x \in M$$

$$c_\mu(a, b) = \{ \langle \mu, a \rangle, \langle \mu, b \rangle \} - \langle \mu, [a, b] \rangle = \langle d\varphi_\mu(a), b \rangle, \quad \forall a, b \in \mathfrak{g}$$



Introduction to Symplectic Geometry

Jean-Louis Koszul
Yiming Zou

$$\begin{aligned} \mu : M &\longrightarrow \mathfrak{g}^* \\ \mu(sx) &= s\mu(x) = \text{Ad}^*(s)\mu(x) + \varphi_\mu(s), \quad \forall s \in G, x \in M \\ c_\mu(a, b) &= \{ \langle \mu, a \rangle, \langle \mu, b \rangle \} - \langle \mu, [a, b] \rangle = \langle d\varphi_\mu(a), b \rangle, \quad \forall a, b \in \mathfrak{g} \end{aligned}$$

Science Press
Beijing

Springer

Volume 1 · Number 1 · 2018

Information Geometry

Editor-in-Chief

Shinto Eguchi (Tokyo)

Co-Editors

Nihat Ay (Leipzig)
Frank Nielsen (Paris)
Jun Zhang (Ann Arbor)

Advisory Board

Ole E. Barndorff-Nielsen (Aarhus)
David Cox (Oxford)
Bradley Efron (Palo Alto)
C.R. Rao (Hyderabad)

Honorary Editors

Shun-ichi Amari (Tokyo)
Imre Csiszár (Budapest)

Associate Editors

Frédéric Barbaresco (Cannes)
Brigo Damiano (London)
Paolo Gibilisco (Rome)
Shiro Ikeda (Tokyo)
Jürgen Jost (Leipzig)
Paul Marriott (Waterloo)
Hiroshi Matsuzoe (Nagoya)
František Matúš (Praha)
Noboru Murata (Tokyo)
Hiroshi Nagaoka (Tokyo)
Jan Naudts (Antwerpen)
Nigel Newton (Colchester)
Richard Nock (Canberra)
Atsumi Ohara (Fukuji)
Giovanni Pistone (Torino)
Constantino Tsallis (Rio de Janeiro)



Springer

Special Issue:

Affine Differential Geometry and Hesse Geometry: A Tribute and Memorial to Jean-Louis Koszul

Submission Deadline: 30th November 2019

Jean-Louis Koszul (January 3, 1921 – January 12, 2018) was a French mathematician with prominent influence to a wide range of mathematical fields. He was a second generation member of Bourbaki, with several notions in geometry and algebra named after him. He made a great contribution to the fundamental theory of Differential Geometry, which is foundation of Information Geometry. The special issue is dedicated to Koszul for the mathematics he developed that bear on information sciences.

Both original contributions and review articles are solicited. Topics include but are not limited to:

- Affine differential geometry over statistical manifolds
- Hessian and Kahler geometry
- Divergence geometry
- Convex geometry and analysis
- Differential geometry over homogeneous and symmetric spaces
- Jordan algebras and graded Lie algebras
- Pre-Lie algebras and their cohomology
- Geometric mechanics and Thermodynamics over homogeneous spaces

Guest Editor:

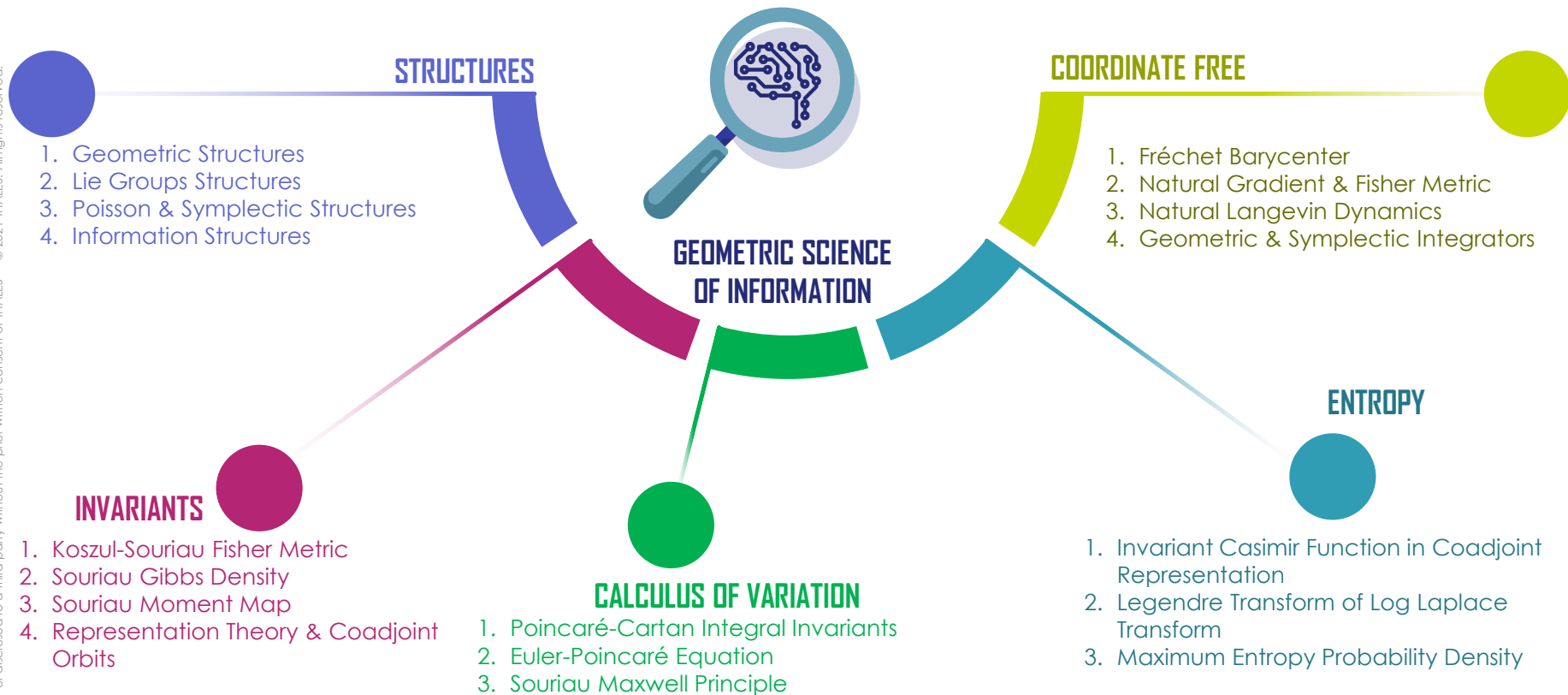
Hideyuki Ishi (Graduate School of Mathematics, Nagoya University)

OPEN

THALES
Building a future we can all trust

Main Concepts behind Symplectic Model of Information Geometry

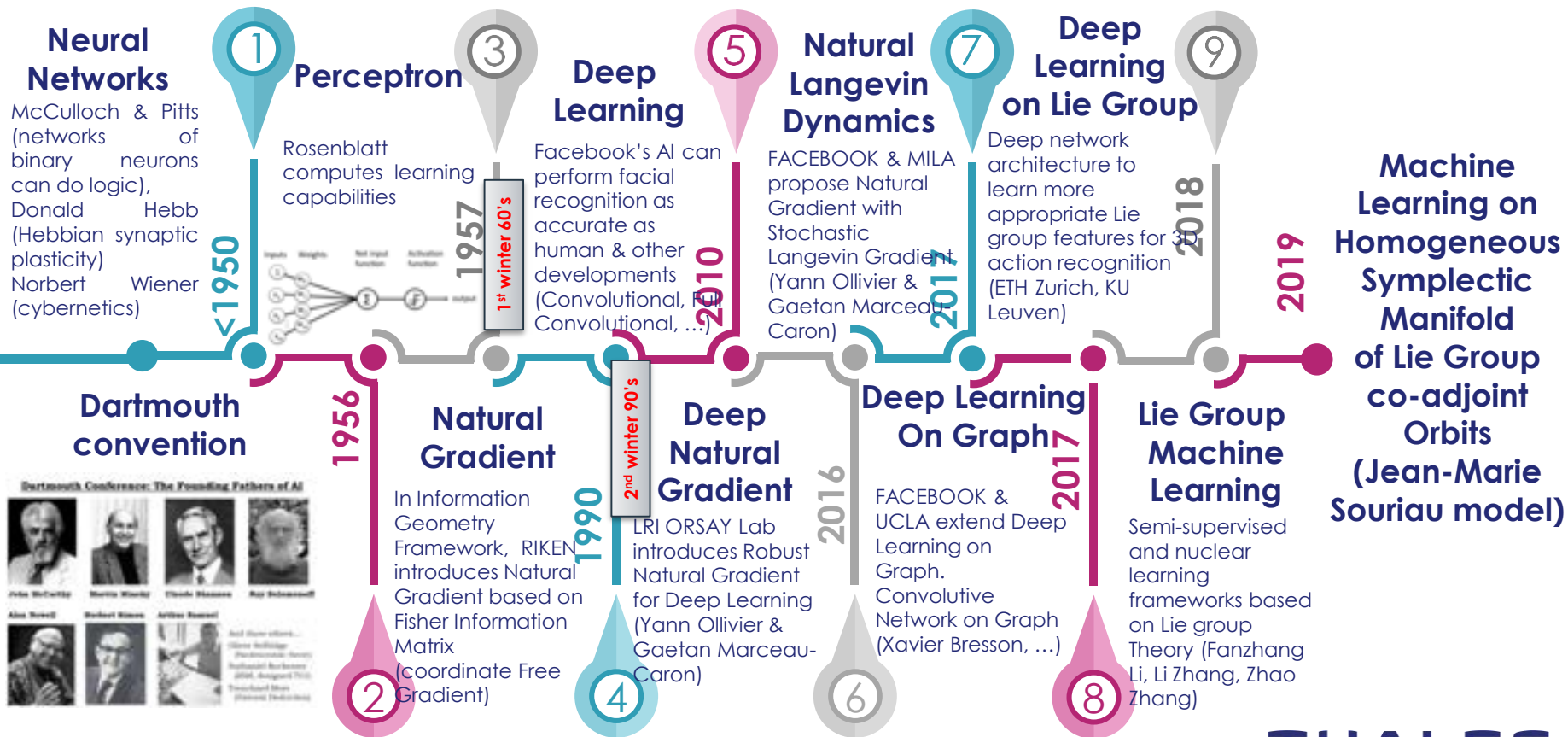
This document may not be reproduced, modified, published, adapted, in any way, in whole or in part or disclosed to a third party without the prior written consent of THALES - © 2021 THALES. All rights reserved.



OPEN

Towards Lie Group & Symplectic Machine Learning

This document may not be reproduced, modified, adapted, published, in whole or in part or disclosed to a third party without the prior written consent of THALES - © 2021 THALES. All rights reserved.



Supervised & Non-Supervised Learning on Lie Groups



Geodesic Natural Gradient on Lie Algebra

Extension of Neural Network Natural Gradient from Information Geometry on Lie Algebra for Lie Groups Machine Learning

Souriau-Fisher Metric on Coadjoint Orbits



Extension of Fisher Metric for Lie Group through homogeneous Symplectic Manifolds on Lie Group Co-Adjoint Orbits

Souriau Exponential Map on Lie Algebra



Exponential Map for Geodesic Natural Gradient on Lie Algebra based on Souriau Algorithm for Matrix Characteristic Polynomial



Souriau Maximum Entropy Density on Co-Adjoint Orbits

Covariant Maximum Entropy Probability Density for Lie Groups defined with Souriau Moment Map, Co-Adjoint Orbits Method & Kirillov Representation Theory

Fréchet Geodesic Barycenter by Hermann Karcher Flow



Extension of Mean/Median on Lie Group by Fréchet Definition of Geodesic Barycenter on Souriau-Fisher Metric Space, solved by Karcher Flow



Symplectic Integrator preserving Moment Map

Extension of Neural Network Natural Gradient to Geometric Integrators as Symplectic integrators that preserve moment map

Mean-Shift on Lie Groups with Souriau-Fisher Distance



Extension of Mean-Shift for Homogeneous Symplectic Manifold and Souriau-Fisher Metric Space



Lie Group Machine Learning

OPEN

LIE GROUP SUPERVISED LEARNING

LIE GROUP NON-SUPERVISED LEARNING

GEOMSTATS: PYTHON Library for Lie Group Machine Learning



hal-02536154, version 1

Pré-publication, Document de travail

Geomstats

<https://github.com/geomstats/geomstats>

Python package 2.1.0 build [coverage](#) [codecov](#) [codecov](#) [codecov](#) [codecov](#) (Coverages for: numpy, tensorflow, pytorch)

Geomstats is an open-source Python package for computations and statistics on manifolds. The package is organized into two main modules: `geometry` and `learning`.

The module `geometry` implements concepts in differential geometry, and the module `learning` implements statistics and learning algorithms for data on manifolds.



To get started with `geomstats`, see the [examples directory](#).

For more in-depth applications of `geomstats`, see the [applications repository](#).

The documentation of `geomstats` can be found on the [documentation website](#).

If you find `geomstats` useful, please kindly cite our [paper](#).

Install geomstats via pip3

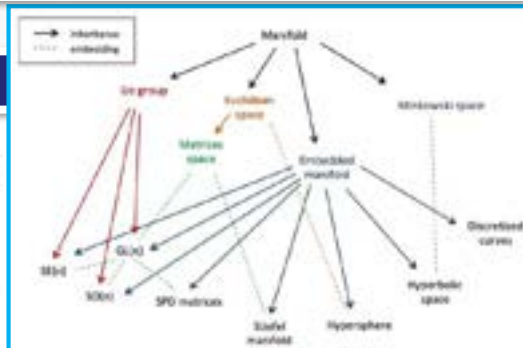
Video: <https://m.youtube.com/watch?v=Ju-Wsd84uG0>

`pip3 install geomstats`

<https://hal.inria.fr/hal-02536154>

Geomstats: A Python Package for Riemannian Geometry in Machine Learning

Nina Miolane¹, Alice Le Brigant², Johan Mathe², Benjamin Hou³, Nicolas Guigui^{4,5}, Yann Thanwerdas^{4,5}, Stefan Heyder⁶, Olivier Peltre, Niklas Koep, Hadi Zaatili⁷, Hatem Hajri⁷, Yann Cabanes, Thomas Gerald, Paul Chauchat⁸, Christian Shewmake, Bernhard Kainz, Claire Donnat⁹, Susan Holmes¹, Xavier Pennec^{4,5} [Détails](#)



PYTHON Library for Machine Learning on Manifold and Lie Group

Computations and statistics on manifolds with geometric structures

- Initiated by INRIA & Stanford University
- Point of contact **Nina Miolane**
(Department of Statistics - Stanford Statistics)
- **PYTHON GEOMSTATS** Package:
 - <https://github.com/geomstats/geomstats>
 - Python Package for Riemannian Geometry in Machine Learning
 - Paper: <https://arxiv.org/abs/1805.08308>

SPECIAL SESSION at GSI'19 conference

- Chaired by **Nina Miolane & Alice Le Brigant**
- www.gsi2019.org

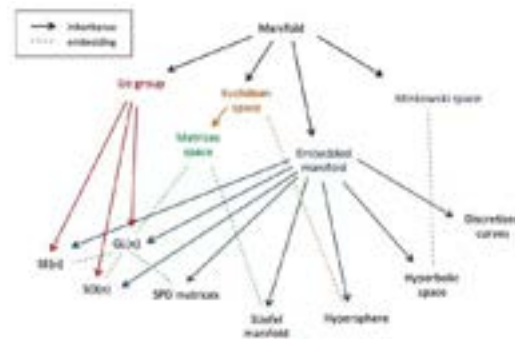
- <https://perso.math.univ-toulouse.fr/statistics-geometry-and-topology/>

SSPD Conference, London, 13th-14th September 2022

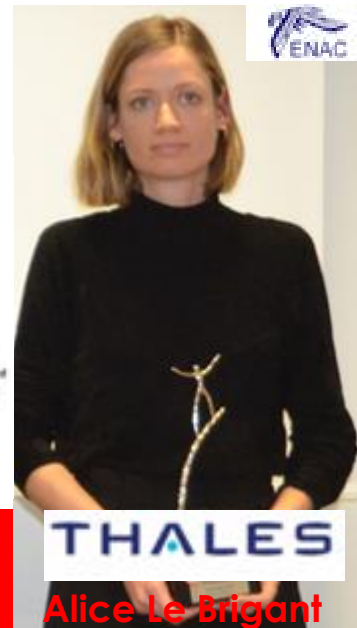
OPEN



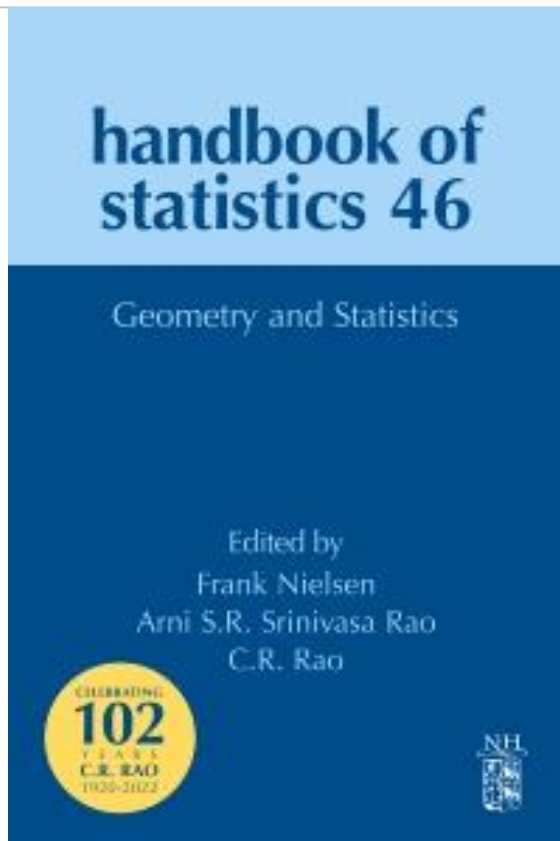
Nina MIOLANE
L'Oréal – Unesco Prize
2016 « Woman in
Science »



Alice LE BRIGANT
THALES PhD Award
2018



Elsevier Handbook of Statistics n°46 « Geometry and Statistics »



Handbook of Statistics

Available online 22 April 2022

In Press, Corrected Proof



Symplectic theory of heat and information geometry

Frédéric Barbaresco

THALES Land & Air Systems, Meudon, France

Available online 22 April 2022.

<https://www.sciencedirect.com/science/article/abs/pii/S0169716122000062>

<https://www.elsevier.com/books/geometry-and-statistics/nielsen/978-0-323-91345-4>

<https://www.sciencedirect.com/handbook/handbook-of-statistics>

World Scientific « **Frontiers in Entropy Across the Disciplines** »

Panorama of Entropy: Theory, Computation, and Applications

Contemporary Mathematics and Its Applications: Monographs, Expositions and Lecture Notes

Frontiers in Entropy Across the Disciplines

Panorama of Entropy: Theory, Computation, and Applications

<https://doi.org/10.1142/12920> | November 2022

Pages: 650

Edited By: Willi Freeden (*University of Kaiserslautern, Germany*) and M Zuhair Nashed (*University of Central Florida, USA*)



Frontiers in Entropy Across the Disciplines

presents a panorama of entropy emphasizing mathematical theory, physical and scientific significance, computational methods, and applications in mathematics, physics, statistics, engineering, biomedical signals, and signal processing.

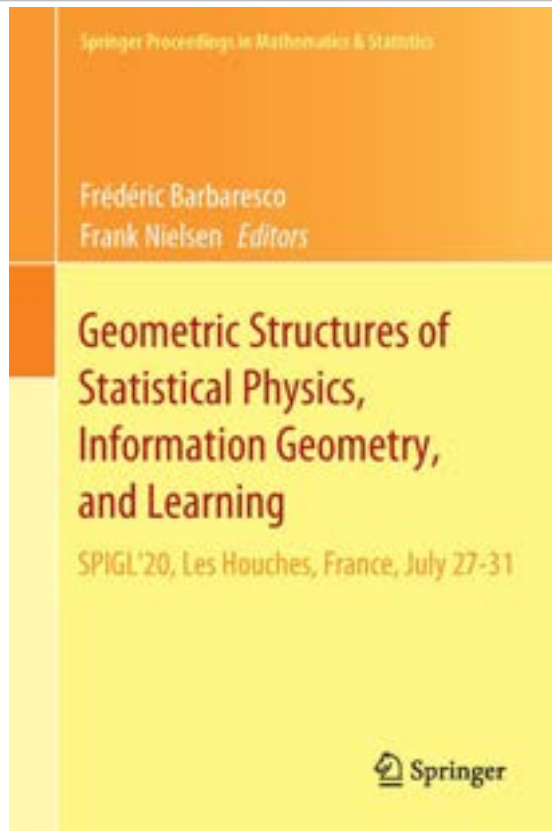
Topics include entropy and society, entropy and time, **Souriau entropy on symplectic model of statistical physics, new definitions of entropy, geometric theory of heat and information**, maximum entropy in Bayesian networks, maximum entropy methods, entropy analysis of biomedical signals (review and comparison of methods), spectral entropy and its application to video coding and speech coding, a comprehensive review of 50 years of entropy in dynamics, a comprehensive review on entropy, entropy-like quantities and applications, topological entropy of multimodal maps, entropy production in complex systems, entropy production and convergence to equilibrium, reversibility and irreversibility in entropy, nonequilibrium entropy, index of various entropy, entropy and the greatest blunder ever.

- **Souriau Entropy based on Symplectic Model of Statistical Physics: Three Jean-Marie Souriau's Seminal Papers on Lie Groups Thermodynamics** (Frédéric Barbaresco)
- **Entropy Geometric Structure as Casimir Invariant Function in Coadjoint Representation: Geometric Theory of Heat and Information Based on Souriau Lie Groups Thermodynamics and Lie Algebra Cohomology** (Frédéric Barbaresco)

<https://www.worldscientific.com/worldscibooks/10.1142/12920#t=aboutBook>

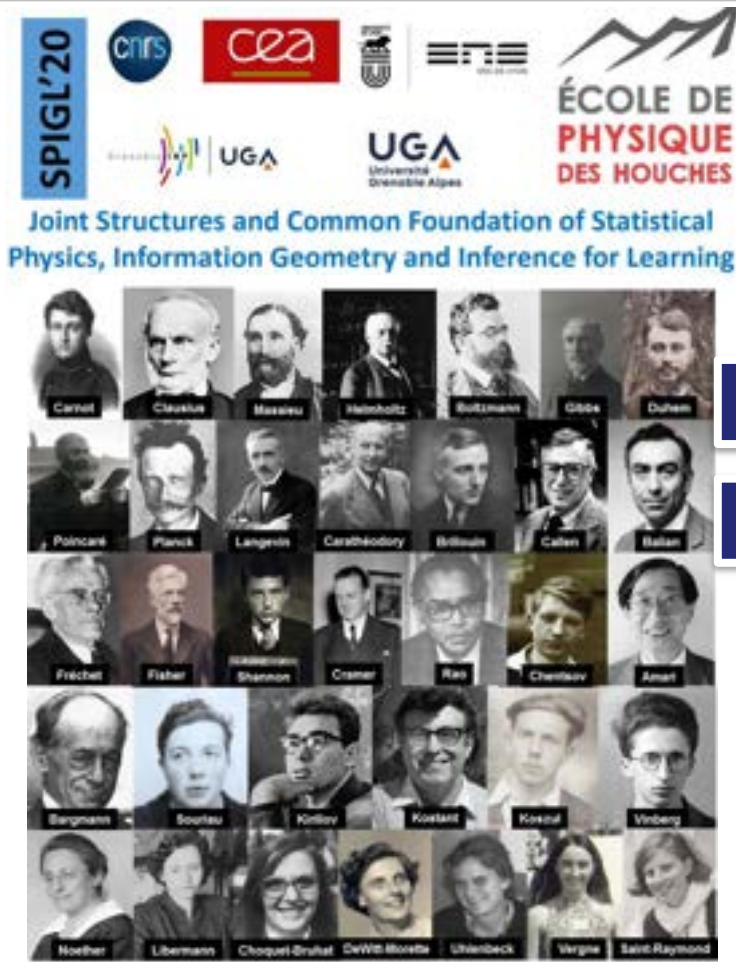
THALES
Building a future we can all trust

Ecole de Physique des Houches SPIGL'20, July 2020



<https://www.springer.com/jp/book/9783030779566>

SSPD Conference, London, 13th-14th September 2022



<https://franknielsen.github.io/SPIG-LesHouches2020/>

<https://www.youtube.com/playlist?list=PLo9ufcrEqwWExTBPgQPJwAJhoUChMbRor>

THALES
Building a future we can all trust



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

41st MaxEnt2022
Conference
JULY 18-22, IHP, PARIS





SEE GSI'23 Geometric Science of Information

Sept. 2023, Saint-Malo, Palais du Grande Large, France 6th Edition and 10th anniversary



Almost 5000 pages in SPRINGER Proceedings



GSI'13

Ecole des Mines de Paris



GSI'15

Ecole Polytechnique



GSI'17

Ecole des Mines de Paris



GSI'19

ENAC, Toulouse



GSI'21

Sorbonne Université



GSI'23

Saint-Malo
Palais du Grand Large

ALES
Building a future we can all trust

SEE GSI'23 Geometric Science of Information

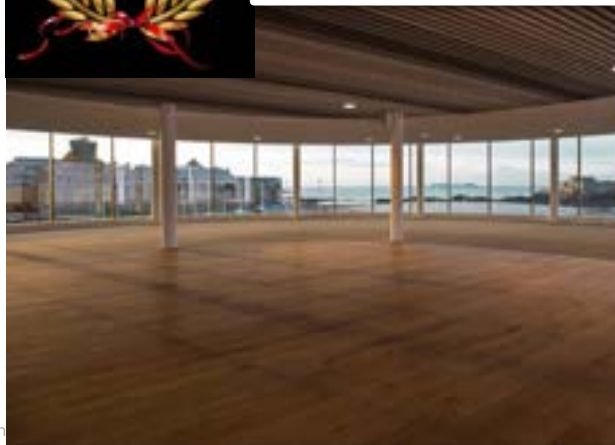
Sept. 2023, Saint-Malo, Palais du Grande Large, France 6th Edition and 10th anniversary



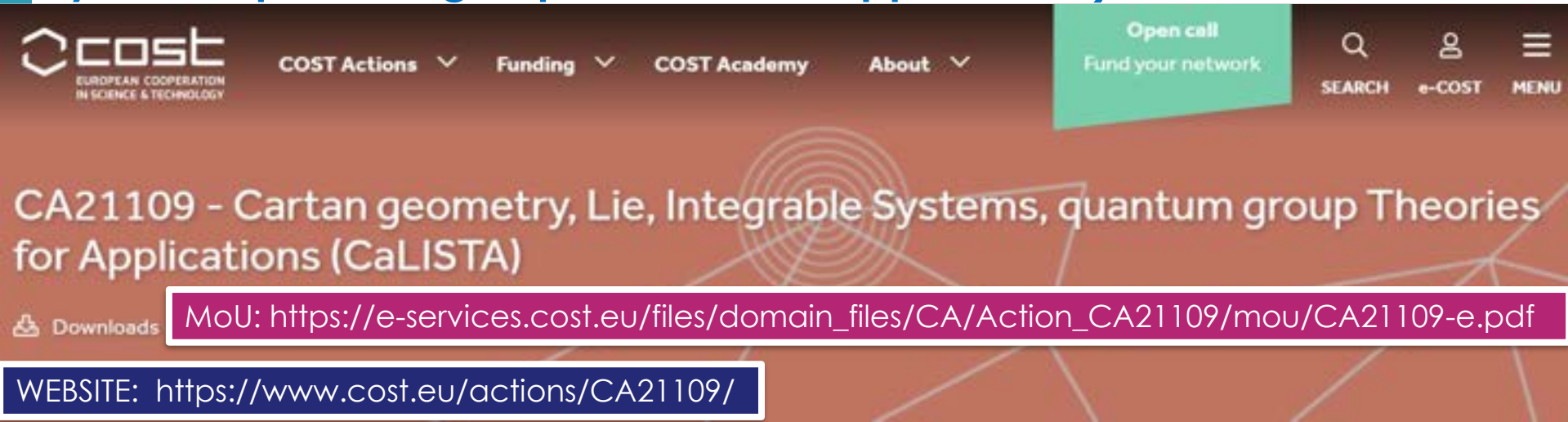
Palais du Grand Large
<https://www.pgl-congres.com/>



SEE GSI'23 « GEOMETRIC SCIENCE OF INFORMATION »



European COST Network CaLISTA (Cartan geometry, Lie, Integrable Systems, quantum group Theories for Applications)



The screenshot shows the top navigation bar of the COST website. On the left is the COST logo with the text 'EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY'. To its right are links for 'COST Actions', 'Funding', 'COST Academy', and 'About'. On the far right is a green button that says 'Open call Fund your network'. Below these are icons for 'SEARCH', 'e-COST', and 'MENU'. The main header area has a large title 'CA21109 - Cartan geometry, Lie, Integrable Systems, quantum group Theories for Applications (CaLISTA)'. Below the title is a purple box containing a download icon and the text 'Downloads'. To the right of this is a white box with a purple border containing the MoU URL: https://e-services.cost.eu/files/domain_files/CA/Action_CA21109/mou/CA21109-e.pdf. At the bottom of the header is a dark blue box with a white border containing the website URL: <https://www.cost.eu/actions/CA21109/>.

MoU: https://e-services.cost.eu/files/domain_files/CA/Action_CA21109/mou/CA21109-e.pdf

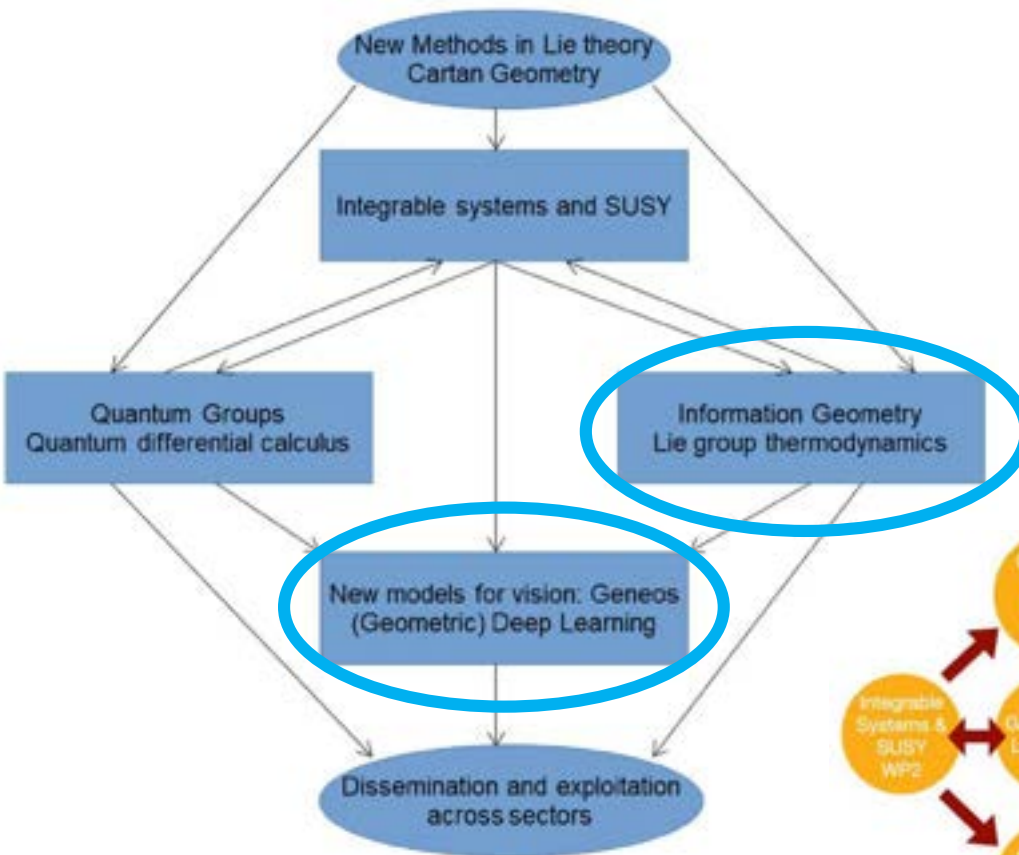
WEBSITE: <https://www.cost.eu/actions/CA21109/>

Description: Symmetry is a central unifying theme in mathematics and physics. This Action focusses on symmetries realized through Lie groups and Lie algebras. In addition to the spectacular achievements in representation theory, and differential geometry, Lie theory is also exceptionally important for the formalization of fundamental physical theories. CaLISTA aims to advance cutting-edge research in mathematics and physics through a systematic application of the ideas and philosophy of Cartan geometry, a thoroughly Lie theoretic approach to differential geometry. In addition to making major progress in Cartan geometry itself, CaLISTA aims to develop crucial applications to integrable systems and supersymmetric gauge theories. Quantum groups and their quantum homogeneous spaces come into the play as a bridge between these topics: quantum groups stem originally from the R-matrix formulation in integrable systems, and their homogeneous spaces offer prototypical examples of noncommutative parabolic geometries. Parabolic geometry is the first and possibly the most important example of Cartan geometry, and one of the main aims of CaLISTA is to obtain a quantum generalization.

Surprisingly, Lie theory and Cartan geometry play a role in an exciting new interpretation of the differential structure, and related dynamics, of models for popular algorithms of vision like Deep Learning and the more recent Geometric Deep Learning. CaLISTA aims to investigate and improve on these techniques. CaLISTA will provide essential mathematical models with far-reaching applications, placing Europe among the leading actors in these innovative research areas.

Action keywords: Lie Theory - Cartan Geometry - Quantum Groups - Integrable Systems - Vision

MARIE SKŁODOWSKA-CURIE ACTIONS **CaLIGOLA**



| | | | | | | | |
|---|---------------------------------------|-----------------|----|------|----|------|-----|
| Work Package no. | 5 | Start/end month | | | | 4-48 | |
| Work Package title | The geometry of Deep Learning | | | | | | |
| Lead beneficiary | Fiorese (UNIBO) [Barbaresco (Thales)] | | | | | | |
| Participating organisation short name | UNIBO | CU | MU | UVEG | UC | ISI | ETH |
| Total person months per participating organisation: | 5 | 9 | 4 | 6 | 6 | 0 | 0 |
| Objectives: Develop new vision models and algorithms with the use of persistent homology, Lie Group statistics to boost the machine learning techniques. Understand (Geometric) Deep Learning with the new invariance mathematical tools and quantum group techniques. | | | | | | | |
| Description of Work and role of specific beneficiaries/associated partners broken down and listed into numbered tasks including the following details: | | | | | | | |
| Task 5.1. Lie groups thermodynamics. Use the techniques of WP1 and WP2 to explore the geometry of the information manifold and its Lie structure (i.e. foliation induced by the coadjoint action). Key expertise: UNIBO, Fiorese, Thales, Barbaresco. | | | | | | | |
| Task 5.2 Persistent Homology and Geneos: going beyond topology and endow the space of all equivariant non-expansive operators (Geneo) a suitable Lie (differentiable) structure. Key expertise: UNIBO, Frosini, UC Otter. | | | | | | | |
| Task 5.3 The geometry of (Geometric) Deep Learning. Differential calculus on graph neural networks, via quantum calculus. New theoretical foundations for the geometric deep learning algorithm. Key expertise: UNIBO (Fiorese, Latini), CU (Ofhuschalla), UC IPAM Institute, Sturmfels. | | | | | | | |



CaLIGOLA: Cartan geometry, Lie and representation theory, Integrable Systems, quantum Groups and quantum computing towards the understanding of the geometry of deep Learning and its Application

References

- [1] F. Barbaresco. Innovative tools for radar signal processing Based on Cartan's geometry of SPD matrices & Information Geometry. IEEE Radar Conf., 2008.
- [2] F. Barbaresco. Robust statistical Radar Processing in Fréchet metric space: OS-HDR-CFAR and OS-STAP Processing in Siegel homogeneous bounded domains. 12th International Radar Symposium (IRS), pages 639–644, 2011.
- [3] F. Barbaresco. Information Geometry of Covariance Matrix: Cartan-Siegel Homogeneous Bounded Domains, Mostow/Berger Fibration and Fréchet Median. Matrix Information Geometry; Springer, p. 199–256, 2012.
- [4] M. Arnaudon, F. Barbaresco, and L. Yang. Riemannian Medians and Means With Applications to Radar Signal Processing. IEEE, August 2013.
- [5] A. Decurninge, F. Barbaresco. Robust Burg Estimation of Radar Scatter Matrix for Mixtures of Gaussian Stationary Autoregressive Vectors. IET Radar, Sonar & Navigation, Volume 11, Issue 1, p. 78–89, January 2016.
- [6] F. Barbaresco. Information Geometry Manifold of Toeplitz Hermitian Positive Definite Covariance Matrices: Mostow/Berger Fibration and Berezin Quantization of Cartan-Siegel Domains. International Journal of Emerging Trends in Signal Processing (IJETSP), March 2013.
- [7] F. Barbaresco and M. Ruiz, "Radar detection for non-stationary Doppler signal in one burst based on information geometry: Distance between paths on covariance matrices manifold," 2015 European Radar Conference (EuRAD), 2015, pp. 41-44
- [8] A. Le Brigant, F. Barbaresco and M. Arnaudon, Geometric barycenters of time/Doppler spectra for the recognition of non-stationary targets, 17th International Radar Symposium, 2016
- [9] Bouleux G., Barbaresco F. (2019) Dilation Operator Approach and Square Root Velocity Transform for Time/Doppler Spectra Characterization on $SU(n)$. In: Nielsen F., Barbaresco F. (eds) Geometric Science of Information. GSI 2019. Lecture Notes in Computer Science, vol 11712. Springer
- [10] B. Balaji and F. Barbaresco "Riemannian mean and space-time adaptive processing using projection and inversion algorithms", Proc. SPIE 8714, Radar Sensor Technology XVII, 31 May 2013
- [11] Y. Cabanes, F. Barbaresco, M. Arnaudon, and J. Bigot. Non-Supervised High Resolution Doppler Machine Learning for Pathological Radar Clutter. IEEE, RADAR 2019, Toulon, France, September 2019.
- [12] F. Barbaresco, Coding & Statistical Characterization of Radar Signal Fluctuation for Lie Group Machine Learning, 2019 International Radar Conference (RADAR), 2019

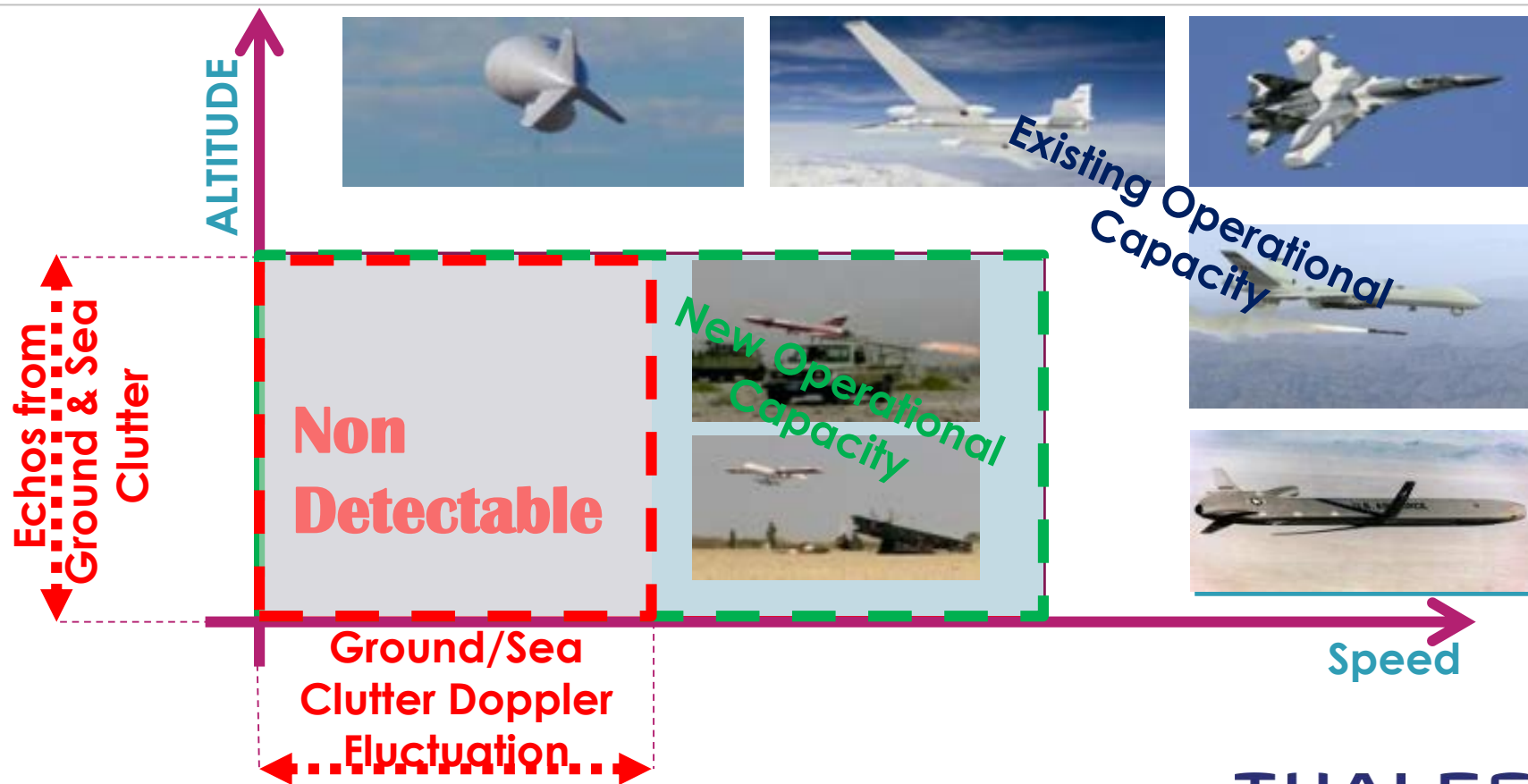
References

- [13] D. A. Brooks, O. Schwander, F. Barbaresco, J. Schneider and M. Cord, Complex-valued neural networks for fully-temporal micro-Doppler classification, 2019 20th International Radar Symposium (IRS), 2019
- [14] D. Brooks, O. Schwander, F. Barbaresco, J.-Y. Schneider and M. Cord, Deep Learning and Information Geometry for Drone Micro-Doppler Radar Classification, 2020 IEEE Radar Conference (RadarConf20), 2020
- [15] P.Y. Lagrave, Y. Cabanes, F. Barbaresco, $SU(1,1)$ Equivariant Neural Networks and Application to Robust Toeplitz Hermitian Positive Definite Matrix Classification. In: Nielsen F., Barbaresco F. (eds) Geometric Science of Information. GSI 2021. LNCS vol 12829. Springer, 2021
- [16] P.Y. Lagrave, F. Barbaresco. Introduction to Robust Machine Learning with Geometric Methods for Defense Applications. 2021. (hal-03309807)
- [17] M. Pilté, S. Bonnabel, and F. Barbaresco. Tracking the frenet-serret frame associated to a highly maneuvering target in 3D. In 2017 IEEE 56th Annual Conference on Decision and Control (CDC), pages 1969–1974, Dec 2017. 9, 23 [18] S. Jouaber, S. Bonnabel, S. Velasco-Forero, M. Pilté. NNAKF: A Neural Network Adapted Kalman Filter for Target Tracking. ICASSP 2021 - 2021 IEEE International Conference ICASSP, Jun 2021
- [19] M. Vincent, A. El Fallah-Seghrouchni, V. Corruble, N. Bernardin, R. Kassab, F. Barbaresco. Monte Carlo Tree Search for Multi-function Radar Task Scheduling. Conference on Artificial Intelligence for Defense, Nov 2021, Rennes, France. (hal-03451838)
- [20] N. Nour, R. Belhaj-Soullami, C. L. R. Buron, A. Peres and F. Barbaresco, Multi-Radar Tracking Optimization for Collaborative Combat, 2021 21st International Radar Symposium (IRS), 2021
- [21] M. Klein et al., AI-Augmented Multi-Function Radar Engineering with Digital Twin: Towards Proactivity, 2020 IEEE Radar Conference (RadarConf20), 2020
- [22] F. Barbaresco, Gaussian Distributions on the Space of Symmetric Positive Definite Matrices from Souriau's Gibbs State for Siegel Domains by Coadjoint Orbit and Moment Map. In: Nielsen F., Barbaresco F. (eds) Geometric Science of Information. GSI 2021. LNCS, vol 12829. Springer
- [23] F. Barbaresco, Koszul lecture related to geometric and analytic mechanics, Souriau's Lie group thermodynamics and information geometry. Information Geometry Journal SPRINGER, 13th Jan. 2021
- [24] F. Barbaresco, Archetypal Model of Entropy by Poisson Cohomology as Invariant Casimir Function in Coadjoint Representation and Geometric Fourier Heat Equation. GSI 2021. vol 12829. Springer, 2021
- [25] F. Barbaresco, Symplectic Theory of Heat and Information Geometry, Handbook of Statistics series, Geometry and Statistics, 1st Edition, Elsevier, July 2022

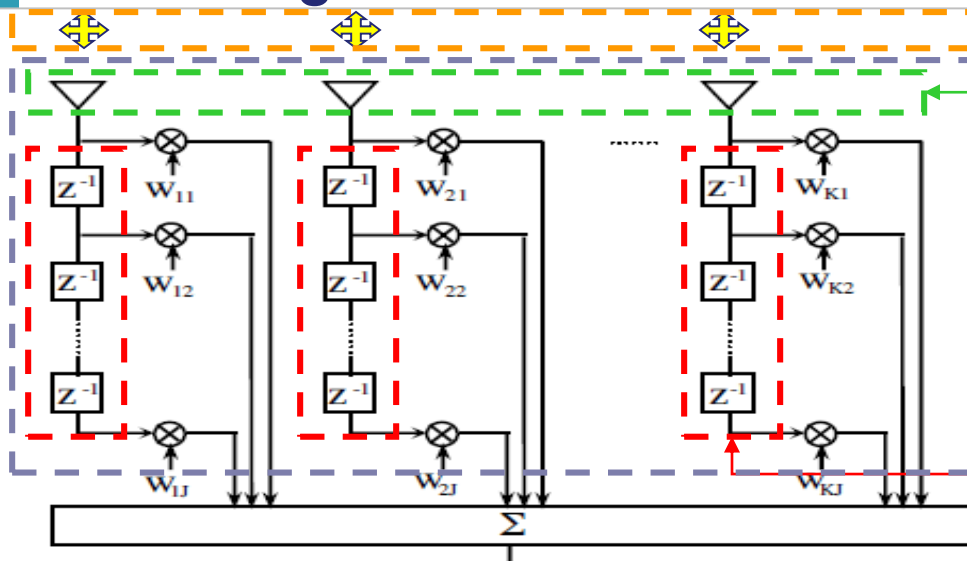
Advanced Sensor Processing based on Koszul Information Geometry

- Geometric Matrix CFAR/STAP for (very) slow targets detection in clutter
- Complex-Valued CNN & Covariance-Matrix-Valued HPDNet for Micro-Doppler ATDR
- Tracker parameters tuning by Deep Learning for tracking hyper-maneuvering targets
- Multi-Agent Reinforcement Learning for Sensor Resources Management for tracking hyper-maneuvering targets
- Multi-Sensors Collaborative Tracking by Distributed Auctions for tracking in saturating scenario (swarm, fleet of targets, ...)

Detection of new threats (slow moving targets)



Spatio-Temporal Digital Measurement (and Polarimetric) of Radar Electromagnetic Wave



Polarimetric Digital Measurement
(Polarimetric Covariance Matrix)

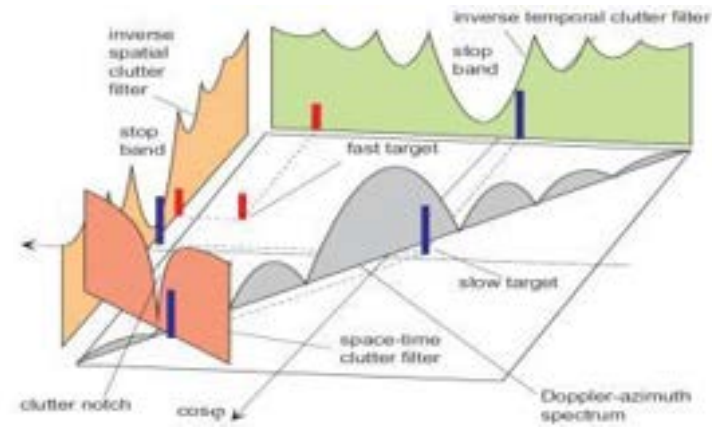
Space-Time Digital Measurement
(Space-Time Covariance Matrix)

Space Digital Measurement
(Space Covariance Matrix)

Time Digital Measurement
(Doppler Covariance Matrix)

Digital Measurement of EM Wave

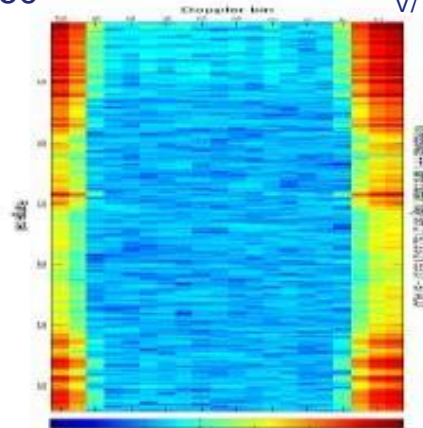
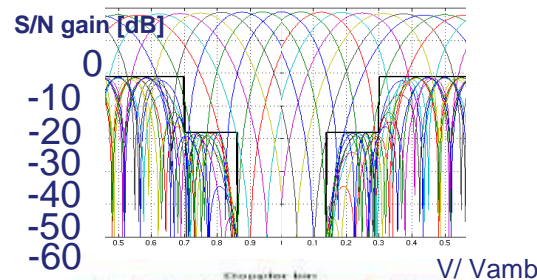
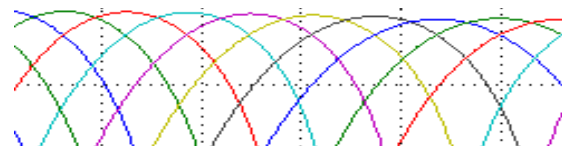
- Polarimetry Measurement => Polar Processing
- Time Measurement => Doppler Processing
- Space Measurement => Antenna Processing
- Space-Time Measurement => STAP Processing



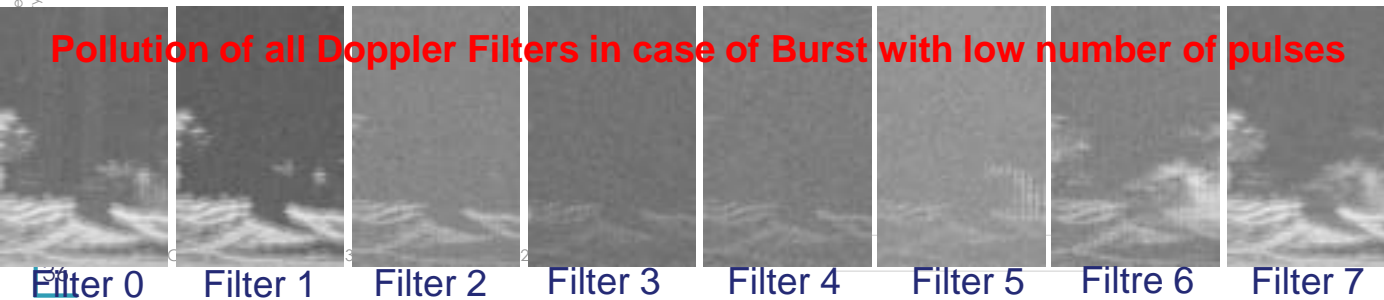
Challenges of Doppler Radar Processing

Detection of objects in Inhomogeneous Ground Clutter

- Classical Doppler Filter Banks (or FFT) are not efficient with very short bursts (<16 pulses):
 - Low Resolution of Doppler Filters with short Bursts (Low sidelobes / high loss, wide filter)
 - If Target Doppler is between two Doppler filters, energy is spread on adjacent filters. Gain between 2 filters is lower than gain at filter center ("Straddling loss")
 - Ground Clutter Energy is not limited to zero-Doppler filter but pollution is spread over all filters due to poor Filter-Banks Resolution & Doppler side lobes in case of very short Bursts.

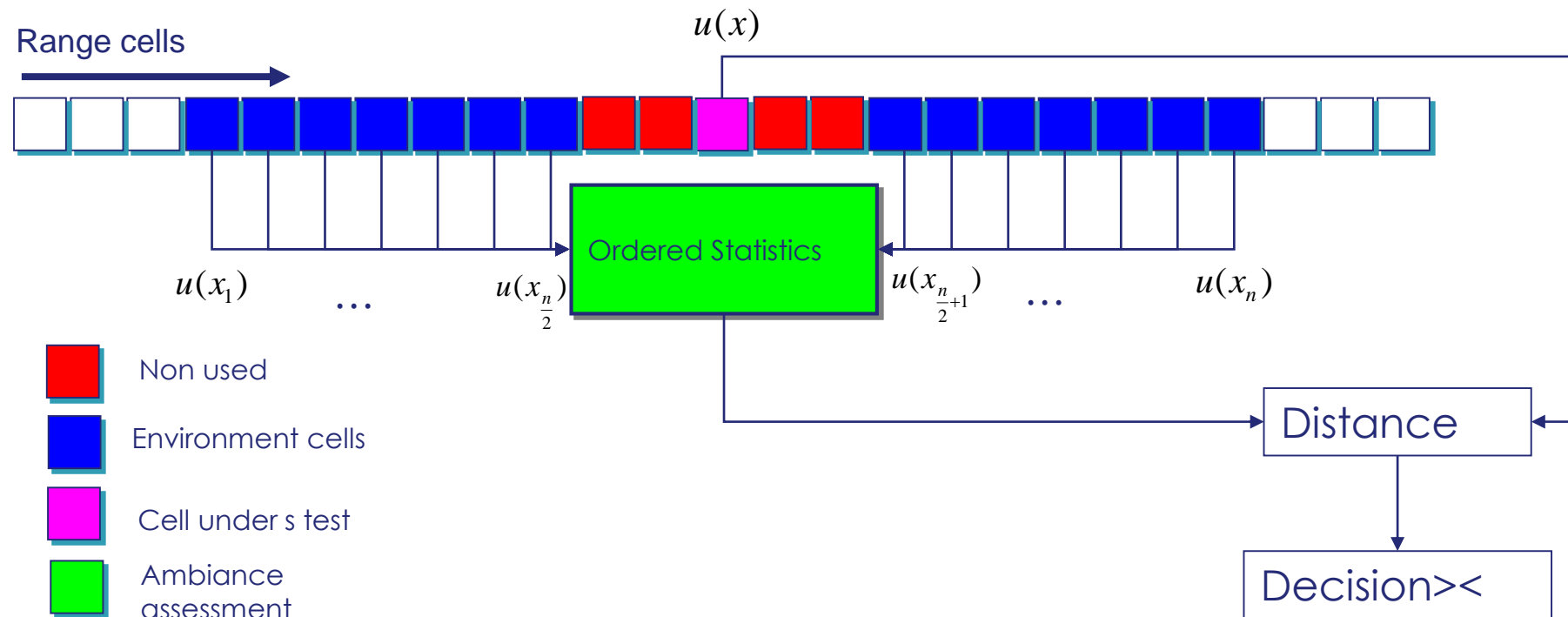


Pollution of all Doppler Filters in case of Burst with low number of pulses



OS-CFAR (Ordered Statistics CFAR)

Classical OS-CFAR on Doppler Filter Bank



Challenges of Doppler Radar Processing: Detection in Sea

Detection of slow targets in Sea Clutter

➤ Sea Clutter is highly inhomogeneous

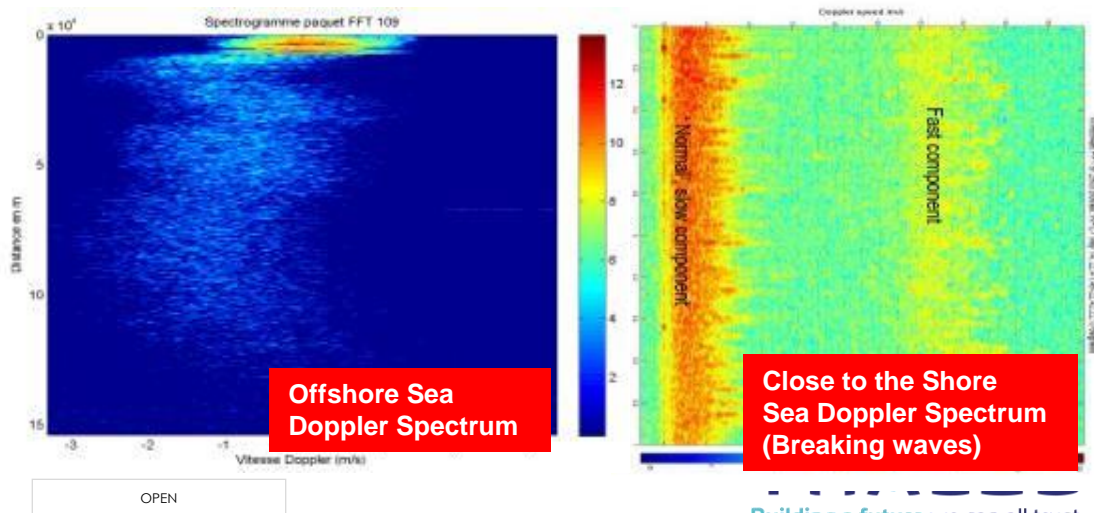
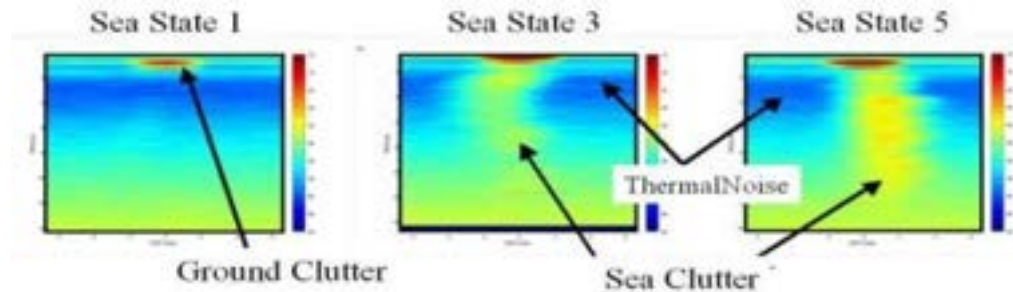
- Doppler fluctuation
- Time/space Fluctuation

➤ Sea Clutter is dependant of

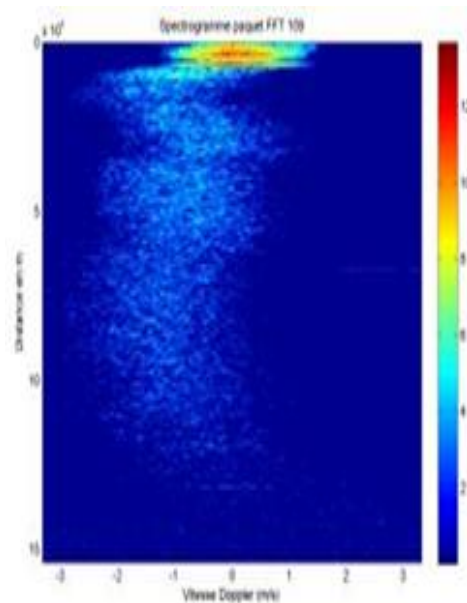
- Sea current
- Surface wind
- fetch
- Bathymetry

➤ Sea Clutter is corrupted by

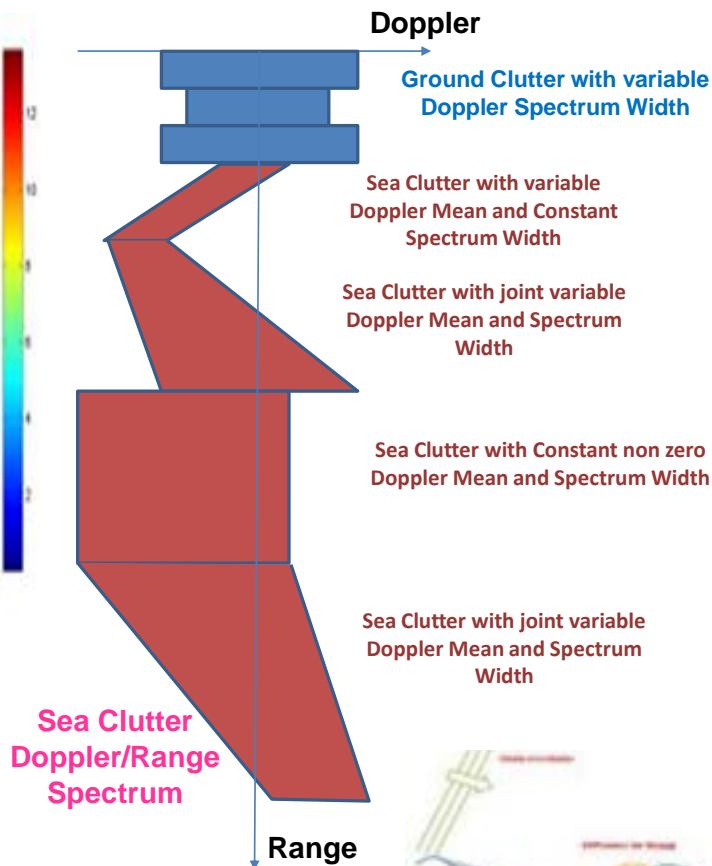
- Spikes due to breaking waves
- "Moutonnement"



Doppler Mean & Doppler Width Variation

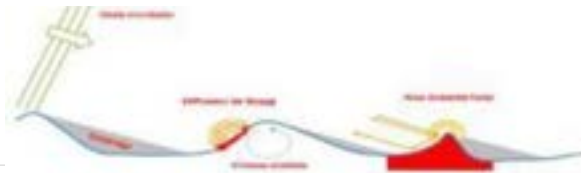


Real Recorder Sea Clutter

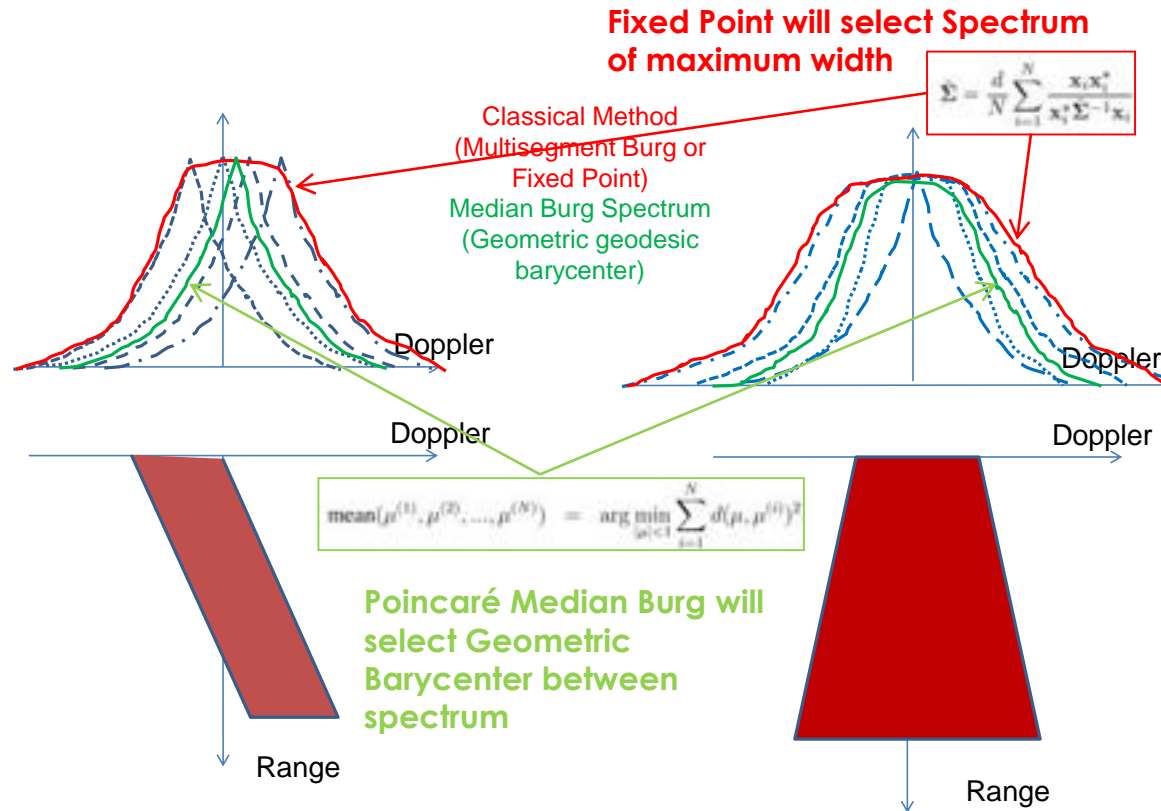


Sea Clutter Variations

- Sea Clutter is highly inhomogeneous
 - Doppler fluctuation
 - Time/space Fluctuation
- Sea Clutter is dependent of
 - Sea current
 - Surface wind
 - fetch
 - Bathymetry
- Sea Clutter is corrupted by
 - Spikes due to breaking waves
 - Mottle ("Moutonement")



Median Estimation of Doppler Spectrum Statistics based on Fréchet Barycenter and Information Geometry versus classical methods

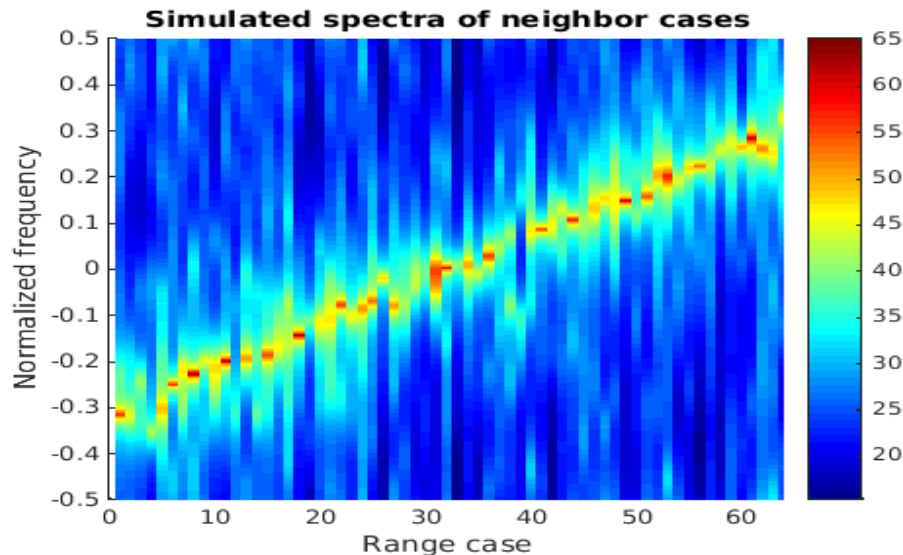
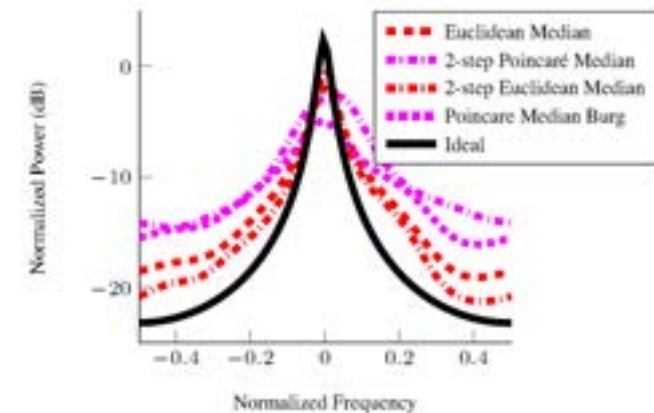
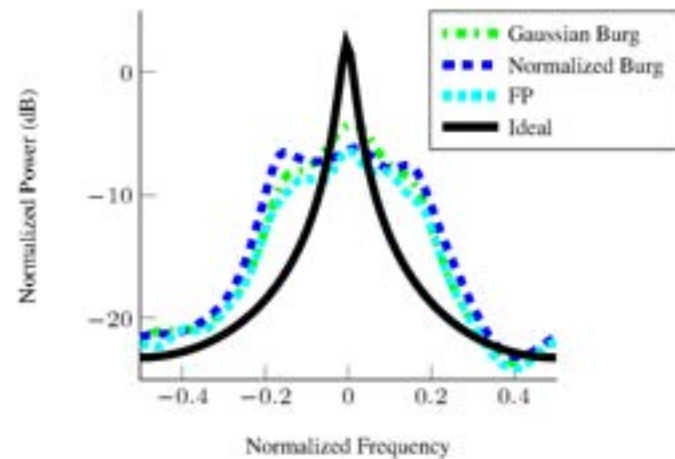


Classical methods for estimation of Mean Clutter Doppler spectrum are based on Multi-segment Burg or “Fixed Point” algorithms using a sliding window along range axis. these approaches suffer of many drawbacks in case of non-stationary clutter.

We propose to estimate Mean Doppler spectrum by Geometric Barycenter

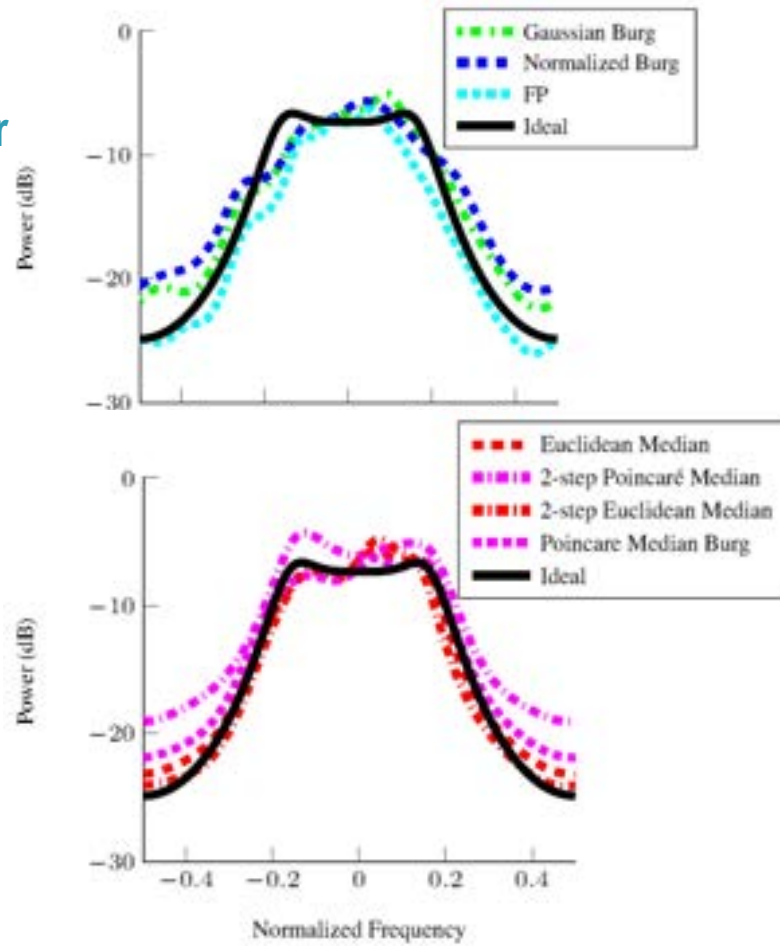
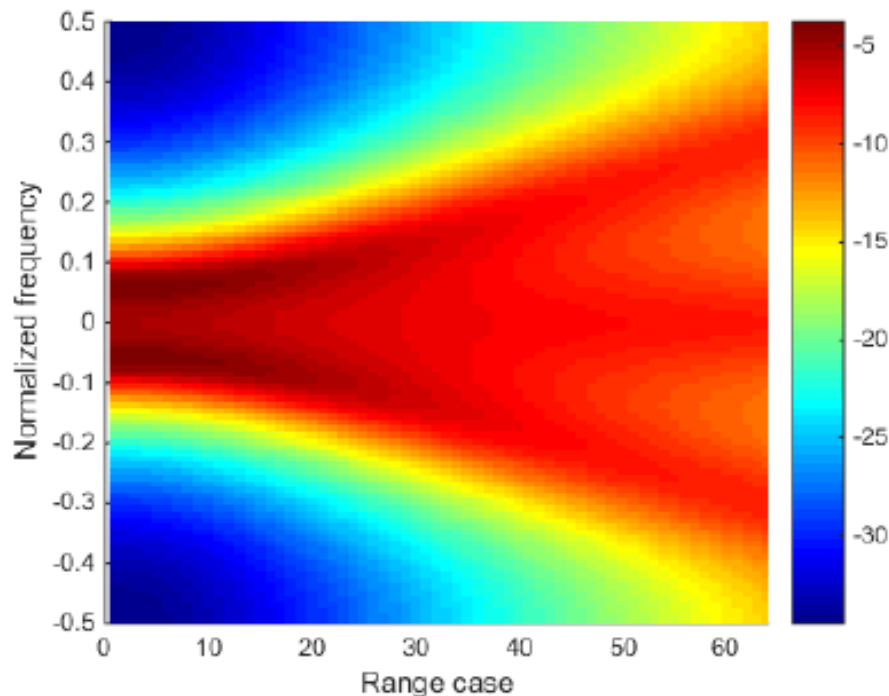
Median Estimation of Doppler Spectrum Statistics based on Fréchet Barycenter and Information Geometry: variation of Doppler Mean

as “Fixed Point” and Multi-segment Burg algorithms take into account all neighbor cases with the same weights, the resulting spectrum is artificially widened. On the contrary, the median-based estimator only depends on the considered Riemannian geometry in the space of covariance matrices, and is able to “interpolate” Doppler spectrum to provide a good estimator of “centrality”.



Median Estimation of Doppler Spectrum Statistics based on Fréchet Barycenter and Information Geometry: variation of Doppler Width

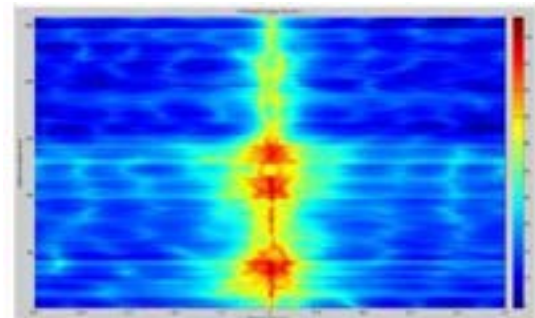
we illustrate the good property of geodesic median L1-barycenter method to estimate Doppler Spectrum in case of non-stationary for clutter Doppler mean



Statistics Random states in a metric space to characterize the fluctuation of the electromagnetic signal

For digital measurement of the electromagnetic wave, main question is about the statistical fluctuations characterization:

- **In amplitude:** variation of the signal in power
- **On polarimetry:** variation of the polarization of the wave
- **Spatially:** variation of the direction of arrival of the wavefront
- **Temporarily:** Doppler spectrum variation



Doppler/Distance Spectrum

Example of the Doppler radar signal relating to the time series of measurements

Measuring the time series of the digitized signal (for a given direction)

$$Z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \quad \text{avec } z_i \in \mathbb{C}$$

$$R_n = E[ZZ^+] = \begin{bmatrix} r_0 & r_1^* & \cdots & r_{n-1}^* \\ r_1 & r_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_1^* \\ r_{n-1} & \cdots & r_1 & r_0 \end{bmatrix}$$

Stationnary \rightarrow R Toeplitz

$$\text{avec } r_k = E[z_m z_{m-k}^*]$$

with r_k : correlation coefficient used to compute the Doppler Spectrum

$$|S_Z(f)|^2 = \left| \sum_{k=-(n-1)}^{n-1} r_k e^{-j2\pi f k} \right|^2$$

Simple case illustration: Toeplitz Hermitian Positive Definite 2x2 matrix

Example with 2x2 Toeplitz Hermitian positive definite covariance matrix:

➤ 1 matrix of covariance of a stationnary 2x2 could be coded by « 1 point » on a product manifold $R_+^* \times D$ (with D Poincaré Unit Disk)

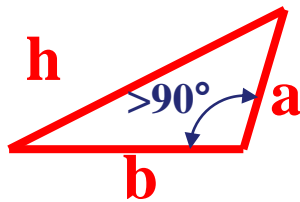
$$\Omega = \begin{bmatrix} h & a - ib \\ a + ib & h \end{bmatrix}$$

with $a, b \in R$ et $h \in R_+^*$

$$\det(\Omega) = h^2 - (a + ib)(a - ib)$$

$$\det(\Omega) = h^2 - (a^2 + b^2) > 0$$

$$h^2 > a^2 + b^2$$



Ambligone triangle
 $h^2 > a^2 + b^2$

$$h^2 > a^2 + b^2$$

$$h^2 = a^2 + b^2$$

$$\Omega^+ = \Omega$$

$$\Omega = h \cdot \begin{bmatrix} 1 & \mu^* \\ \mu & 1 \end{bmatrix} \quad \text{with} \quad \mu = \frac{a + ib}{h}$$

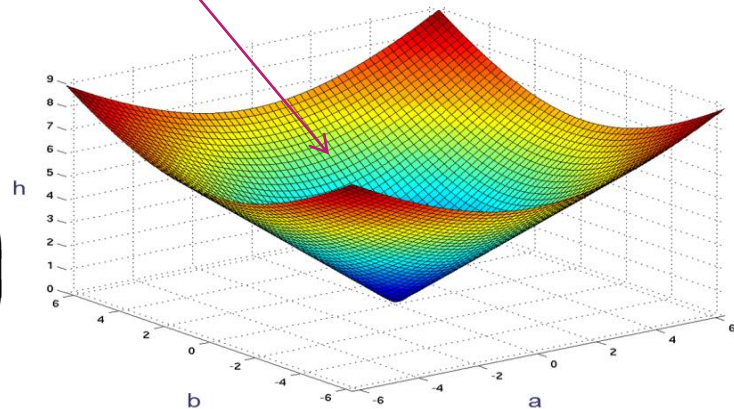
$$|\mu|^2 = \frac{a^2 + b^2}{h^2} < 1 \quad \leftarrow h^2 > a^2 + b^2$$

$$\mu \in D = \{z / |z| < 1\}$$

Doppler Mean

$\phi_{Doppler}$: phase of μ

$$\phi_{Doppler} = \arctan\left(\frac{b}{a}\right)$$



OPEN

What are the constraints between the parameters of the covariance matrix of a stationary (circular) complex signal

■ The parameters of a covariance matrix of a time series of a stationary (circular) complex signal are constrained by:

➤ Toeplitz Structure (same elements on diagonals) :

$$\forall n, \quad E[z_n z_{n-k}^*] = r_k$$

$$R_n = \begin{bmatrix} r_0 & r_1^* & r_2^* & \cdots & r_{n-1}^* \\ r_1 & r_0 & r_1^* & \ddots & \vdots \\ r_2 & r_1 & \ddots & \ddots & r_2^* \\ \vdots & \ddots & \ddots & r_0 & r_1^* \\ r_{n-1} & \cdots & r_2 & r_1 & r_0 \end{bmatrix} = \left[\begin{array}{c|c} R_{n-1} & \begin{bmatrix} r_{n-1}^* \\ \vdots \\ r_2^* \\ r_1^* \end{bmatrix} \\ \hline r_{n-1} & \cdots & r_2 & r_1 & r_0 \end{array} \right]$$



Otto Toeplitz

➤ Hermitian Structure:

$$R_n^+ = R_n \quad \text{With + transposed \& Conjugated}$$

➤ Positive Definite Structure (positive eigenvalues) :

$$\forall Z \in C^n, \quad Z^+ R_n Z > 0 \quad \text{and} \quad \lambda_i > 0 \quad i = 1, \dots, n \quad \text{with} \quad \det(R_n - \lambda I) > 0$$

Trench/Verblunsky Theorem (Partial Iwasawa Decomposition) Toeplitz Hermitian Positive Definite Matrix

Structure: Toeplitz Hermitian Positive Definite THDP

- All THDP matrices are diffeomorphic to the Product space : $(P_0, \mu_1, \dots, \mu_n) \in \mathbb{R}^+ \times D_n$:
- P_0 is a scale parameter (Radar signal power)
 - μ_k are "shape" parameters (Shape of Spectrum) called réflexion/Verblunsky coefficients

$$\begin{aligned} \varphi : THDP(n) &\rightarrow \mathbb{R}_+^* \times D^{n-1} \\ R_n &\mapsto (P_0, \mu_1, \dots, \mu_n) \\ \text{with } D &= \{z \in \mathbb{C} / |z| < 1\} \end{aligned}$$

- The following Block decomposition is related to Partial Iwasawa decomposition:

$$R_n^{-1} = \begin{bmatrix} \alpha_{n-1} & \alpha_{n-1} A_{n-1}^+ \\ \alpha_{n-1} A_{n-1} & R_{n-1}^{-1} + \alpha_{n-1} A_{n-1} A_{n-1}^+ \end{bmatrix}$$

$$\alpha_n^{-1} = \left[1 - |\mu_n|^2 \right] \alpha_{n-1}^{-1} \quad P_0 = \alpha_0^{-1}$$

$$A_n = \begin{bmatrix} A_{n-1} \\ 0 \end{bmatrix} + \mu_n \begin{bmatrix} A_{n-1}^{(-)} \\ 1 \end{bmatrix}$$

$$V^{(-)} = J.V^*$$

Trench/Verblunsky Theorem and Cholesky Decomposition

Block Structure build iteratively:

$$R_n^{-1} = \begin{bmatrix} \alpha_{n-1} & \alpha_{n-1} A_{n-1}^+ \\ \alpha_{n-1} A_{n-1} & R_{n-1}^{-1} + \alpha_{n-1} A_{n-1} A_{n-1}^+ \end{bmatrix}$$

Generation of (André-Louis) Cholesky Decomposition

$$\Omega_n = (\alpha_n R_n)^{-1} = \Omega_n^{1/2} \cdot \Omega_n^{1/2+}$$

$$\Omega_n^{1/2} = \sqrt{1 - |\mu_n|^2} \begin{bmatrix} 1 & 0_{n-1}^+ \\ A_{n-1} & \Omega_{n-1}^{1/2} \end{bmatrix}$$

A. Cholesky, Sur la résolution numérique des systèmes d'équations linéaires, Manuscrit, Fonds A. Cholesky, Archives de l'Ecole Polytechnique, Palaiseau.



A.L. Cholesky
(Cholesky Artillery officer killed during
1st World war)



Computing
MachineDact
yle

THALES
Building a future we can all trust

Fisher Metric & Entropy Hessian Metric of “Information Geometry”

We consider as metric, **hessian of Entropy**:

- Entropy: $S(\bar{R}_n) = \log(\det \bar{R}_n^{-1}) - \log(\pi.e)$
- By using Block structure, and determinant lemma given by:

$$\det \begin{bmatrix} p & q^+ \\ q & Q \end{bmatrix} = p \cdot \det(Q - qq^+)$$

$$\bar{R}_n^{-1} = \begin{bmatrix} \bar{\alpha}_{n-1} & \bar{\alpha}_{n-1} \bar{A}_{n-1}^+ \\ \bar{\alpha}_{n-1} \bar{A}_{n-1} & \bar{R}_{n-1}^{-1} + \bar{\alpha}_{n-1} \bar{A}_{n-1} \bar{A}_{n-1}^+ \end{bmatrix} \quad \begin{aligned} \bar{\alpha}_n^{-1} &= [1 - |\bar{\mu}_n|^2] \bar{\alpha}_{n-1}^{-1} \\ \bar{P}_0 &= \bar{\alpha}_0^{-1} \end{aligned}$$

➔
$$S(\bar{R}_n) = - \sum_{k=1}^{n-1} (n-k) \log[1 - |\bar{\mu}_k|^2] - n \log[\pi.e.\bar{P}_0]$$

$$V = \sum_{i=1}^n k_i \log(1 - x_i \bar{x}_i) \quad (k_i \text{ Konstanten})$$

Example of 1942 E. Kähler paper hyper-abelian case)



E. Kähler

Entropy is parameterized by Estimators of Verblunski parameters:

$$\theta^{(n)} = [\bar{P}_0 \quad \bar{\mu}_1 \quad \cdots \quad \bar{\mu}_{n-1}]^T = E[[P_0 \quad \mu_1 \quad \cdots \quad \mu_{n-1}]^T]$$

Information Geometry provide a natural metric given by the Hessian of the Entropy:

$$g_{ij} = \frac{\partial^2 S}{\partial \theta_i \partial \theta_j^*}$$

$$ds_{dual}^2 = \sum_{i,j} g_{ij}^{dual} d\bar{\theta}_i^{(n)} d\bar{\theta}_j^{(n)*} = n \cdot \left(\frac{d\bar{P}_0}{\bar{P}_0} \right)^2 + \sum_{i=1}^{n-1} (n-i) \frac{|d\bar{\mu}_i|^2}{(1 - |\bar{\mu}_i|^2)^2}$$

Robust Reflection coefficients estimation by Regularized Burg Algo.

Complex autoregressive parameters $A_n = [a_1^{(n)} \quad \dots \quad a_n^{(n)}]^T$ **and reflections coefficients** $\{\mu_i\}_{i=1}^{N-1}$ **are computed by Regularized Burg algorithm from pulses of each radar burst:** $\{z(k)\}_{k=1}^N$

$$f_0(k) = b_0(k) = z(k), \quad k=1, \dots, N, \quad \bar{P}_0 = \frac{1}{N} \cdot \sum_{k=1}^N |z(k)|^2 \quad \text{and} \quad a_0^{(0)} = 1$$

For $n = 1$ to $N-1$

$$\bar{\mu}_n = - \frac{\frac{2}{N-n} \sum_{k=n+1}^N f_{n-1}(k) \cdot b_{n-1}^*(k-1) + 2 \cdot \sum_{k=1}^{n-1} \beta_k^{(n)} \cdot a_k^{(n-1)} \cdot a_{n-k}^{(n-1)}}{\frac{1}{N-n} \sum_{k=n+1}^N |f_{n-1}(k)|^2 + |b_{n-1}(k-1)|^2 + 2 \cdot \sum_{k=0}^{n-1} \beta_k^{(n)} \cdot |a_k^{(n-1)}|^2}$$

with $\beta_k^{(n)} = \gamma(2\pi)^2(k-n)^2$

For $k=1$ to $n-1$

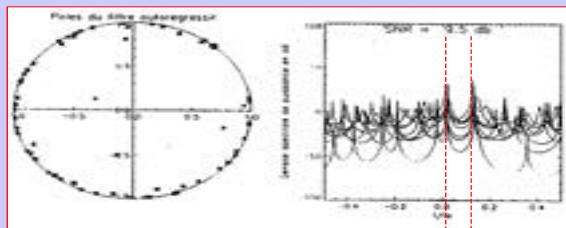
$$\begin{cases} a_0^{(n)} = 1 \\ a_k^{(n)} = a_k^{(n-1)} + \bar{\mu}_n \cdot a_{n-k}^{(n-1)*} \\ a_n^{(n)} = \bar{\mu}_n \end{cases}, \quad \begin{cases} f_n(k) = f_{n-1}(k) + \bar{\mu}_n \cdot b_{n-1}(k-1) \\ b_n(k) = b_{n-1}(k-1) + \bar{\mu}_n^* \cdot f_{n-1}(k) \end{cases}$$



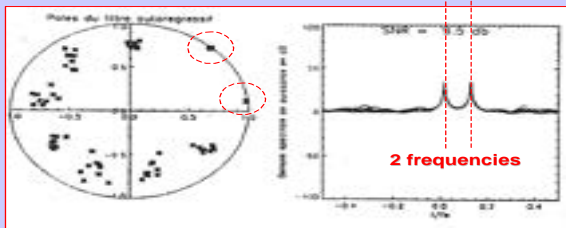
John Parker Burg
(Stanford University)

Regularized Burg Algorithm
THALES PATENT ©

Regularized Burg Algorithm Properties

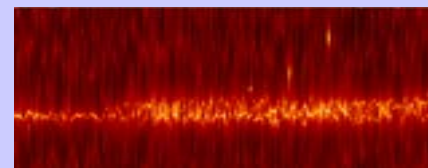
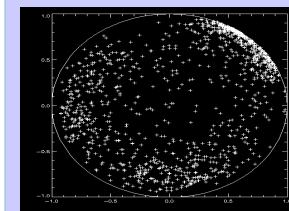
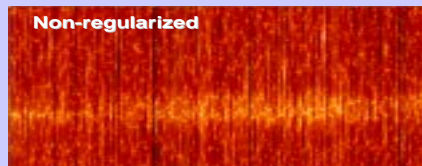


**Non-regularised
Burg Algorithm
(model order : 9)**

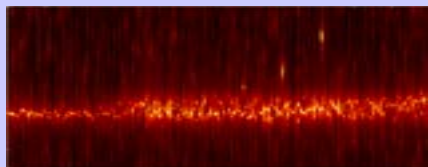
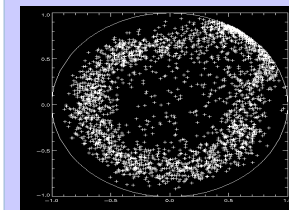


**Regularised
Burg Algorithm
(model order : 9)**

2 frequencies

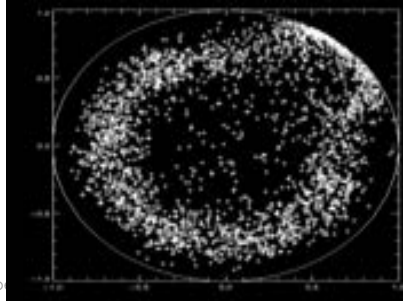
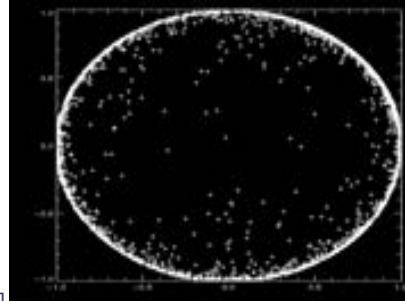
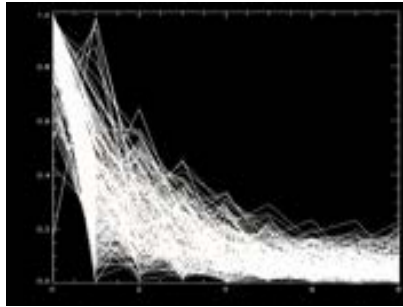
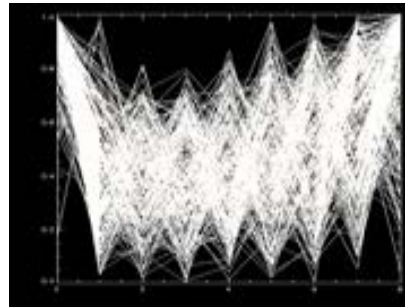
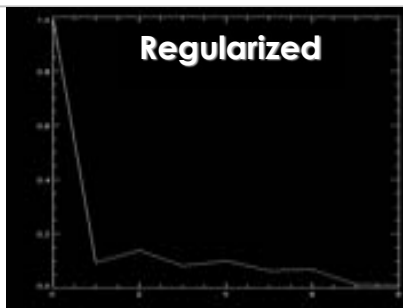
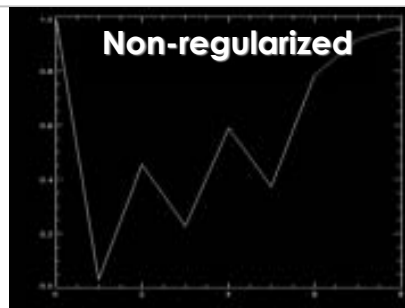


Truncated Model Order



Régularized of Max Order

REGULARIZED BURG ALGORITHM RESULTS



$$A = [X^+ . X + \gamma . \Delta_k]^{-1} . X^+ . x$$

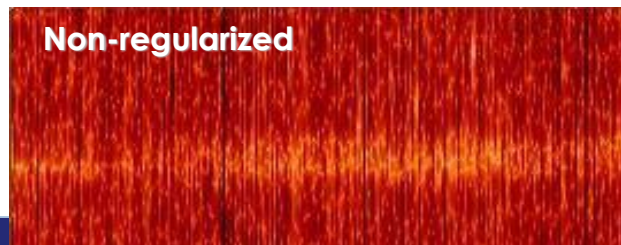
$$\text{for } \gamma = \infty, \quad A \approx \frac{1}{\gamma} \Delta_k . X^+ . x = \frac{\delta}{\gamma}$$

$$\text{then } A^{(n)}(z) \approx z^n + \sum_{k=1}^n \frac{\delta_k}{\gamma} z^{n-k} = Q^{(n)}(z)$$

$$\text{Set } y = \gamma^{1/n} . z \quad \text{et } \varepsilon = \gamma^{-1/n} \Rightarrow Q_\varepsilon(y) = y^n + \sum_{k=1}^n \delta_k . \varepsilon^{n-k} . y^{n-k}$$

$$\lim_{\varepsilon \rightarrow 0} Q_\varepsilon(y) = y^n + \delta_n \Rightarrow r_k^0 = \rho^{1/n} . e^{jk\phi/n} \quad \text{with } \delta_n = \rho . e^{j\phi}$$

$$\Rightarrow z_n \approx \frac{\rho^{1/n} . e^{jk\phi/n}}{\gamma^{1/n}} \quad \text{for } \gamma = \infty$$

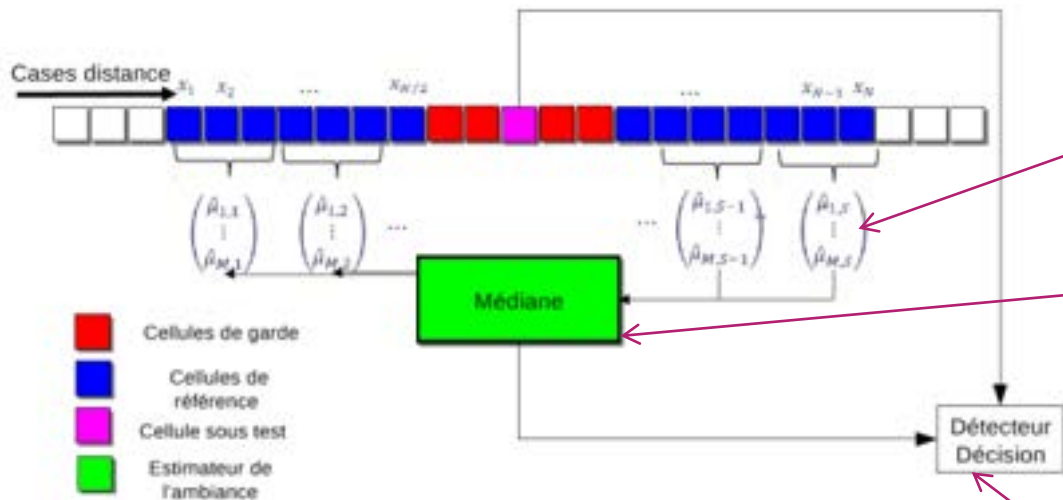


OPEN

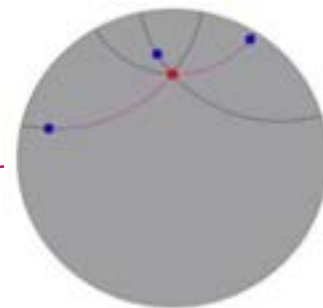
Normalized Median Burg Algorithm

Normalized Median Burg Algorithm

- Normalized Burg estimation on 3 cases
- Median of normalized reflection coefficients
- Normalized GLR detector



$$\hat{\mu}_{m+1} = f^{-1}\left(\left|\hat{\mu}_{m+1}^{(b)}\right|\right) \frac{\hat{\mu}_{m+1}^{(b)}}{\left|\hat{\mu}_{m+1}^{(b)}\right|}$$



$$GLRT(z) = \max_{\theta \in [-0.5; 0.5[} \frac{|p(\theta) + \hat{\Sigma}^{-1}z|^2}{(z + \hat{\Sigma}^{-1}z)(p(\theta) + \hat{\Sigma}^{-1}p(\theta))}$$

Maurice Fréchet : Geodesic Barycenter in Metric Space

This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of THALES - © 2021 THALES. All rights reserved.



Fréchet 1943 Seminal Paper (Clairaut Equation)

Fréchet, M. Sur l'extension de certaines évaluations statistiques au cas de petits échantillons. Revue de l'Institut International de Statistique 1943, 11, 182–205.

Etude des densités distinguées. Appelons (provisoirement, dans ce mémoire) *densité distinguée*, toute densité de probabilité $f(x, \theta)$ telle que la fonction

$$(46) \quad \theta + \frac{\frac{\partial L f(x, \theta)}{\partial \theta}}{\int_{-\infty}^{+\infty} \left[\frac{\partial}{\partial \theta} f(x, \theta) \right]^2 \frac{dx}{f(x, \theta)}}$$

soit indépendante de θ .

Pour ces densités distinguées, on va pouvoir déterminer la fonction minimisante $H'(X_1, \dots, X_n)$ et étendre au cas des petits échantillons la comparaison des méthodes d'estimation faites par divers auteurs dans le cas des grands échantillons. Il vaut donc la peine de chercher la forme générale de $f(x, \theta)$ pour cette catégorie de variables.

de θ . En appelant $h(x)$ cette fonction, on voit qu'on a l'identité de la forme

$$(47) \quad \lambda(\theta) \frac{\partial}{\partial \theta} L f(x, \theta) = h(x) - \theta$$

où $\lambda(\theta) > 0$. On peut considérer $\frac{1}{\lambda(\theta)}$ comme la dérivée seconde d'une fonction $\mu(\theta)$; d'où $\frac{\partial}{\partial \theta} L f(x, \theta) = \mu_{\theta}''(\theta) [h(x) - \theta]$.

Par suite $L f(x, \theta) - \mu_{\theta}'(\theta) [h(x) - \theta] - \mu(\theta)$ est une quantité indépendante de θ que nous pouvons représenter par $l(x)$.

Ainsi toute densité distinguée, $f(x, \theta)$, est de la forme

$$(48) \quad f(x, \theta) = e^{\mu_{\theta}'(\theta) [h(x) - \theta] + \mu(\theta) + l(x)}$$

$$(52\text{bis}) \quad \lambda \mu'' = 1.$$

Incidemment, puisque, d'après (52), $\lambda(\theta)$ est positif, il en résulte aussi que $\mu'' \left(= \frac{1}{\lambda(\theta)} \right)$ est aussi positif. **Fisher metric**

On peut d'ailleurs préciser d'une manière plus directe que par (50), le choix des fonctions $\mu(\theta)$, $h(x)$, $l(x)$: on peut prendre arbitrairement $h(x)$ et $l(x)$ et alors $\mu(\theta)$ est déterminé par (50) ou même mieux par une formule explicite. En effet, (50) peut s'écrire

$$e^{\theta \mu' - \mu} = \int_{-\infty}^{+\infty} e^{s' h(x) + l(x)} dx.$$

Donnons-nous alors arbitrairement $h(x)$ et $l(x)$ et soit s une variable arbitraire: la fonction

$$\int_{-\infty}^{+\infty} e^{s h(x) + l(x)} dx \quad (1)$$

sera une fonction positive connue que nous pourrions représenter par $e^{\psi(s)}$. On voit alors que $\mu(\theta)$ sera défini par

$$\theta \mu' - \mu = \psi(\mu')$$

$$(55) \quad \mu = \theta \mu' - \psi(\mu') \quad \text{Legendre-Clairaut}$$

c'est-à-dire une équation de Clairaut. La solution $\mu' = \text{constante}$ réduirait $f(x, \theta)$, d'après (48) à une fonction indépendante de θ , cas où le problème n'aurait plus de sens. μ est donc donné par la solution singulière de (55), qui est unique et s'obtient en éliminant s entre $\mu = \theta s - \psi(s)$ et $\theta = \psi'(s)$ ou encore entre

$$e^{\theta \mu' - \mu} = \int_{-\infty}^{+\infty} e^{s h(x) + l(x)} dx \text{ et}$$

$$(55\text{bis}) \quad \int_{-\infty}^{+\infty} e^{s h(x) + l(x)} [h(x) - \theta] dx = 0.$$

Si l'on veut, $\mu(\theta)$ est donné par la relation

$$e^{-\mu} = e^{-\theta s} \int_{-\infty}^{+\infty} e^{s h(x) + l(x)} dx$$

où s est donné en fonction de θ par la relation implicite (55bis).

Mean of structured covariance matrices: Fréchet Barycenter

Distance between 2 covariance matrices is given in metric space of Product Manifold $R^+ \times D_n$ (via Poincaré distance in the Unit Disk):

$$d^2 \left[\left(\bar{P}_{0,1}, \{\bar{\mu}_{i,1}\}_{i=1}^{N-1} \right), \left(\bar{P}_{0,2}, \{\bar{\mu}_{i,2}\}_{i=1}^{N-1} \right) \right] = N \log^2 \left(\frac{\bar{P}_{0,2}}{\bar{P}_{0,1}} \right) + \sum_{i=1}^{N-1} (N-i) \left(\frac{1}{2} \log \left(\frac{1+\delta_i}{1-\delta_i} \right) \right)^2$$

with $\delta_i = \left| \frac{\bar{\mu}_{i,1} - \bar{\mu}_{i,2}}{1 - \bar{\mu}_{i,1} \bar{\mu}_{i,2}^*} \right|$

The mean is then defined as the geodesic cFréchet centroid which minimises the sum of the squares of the geodesic distances:

$$\left(P_{0,barycenter}, \{\mu_{i,barycenter}\}_{i=1}^{N-1} \right) =$$

$$\underset{P_{0,median}, \{\mu_{i,median}\}_{i=1}^{N-1}}{\text{ArgMin}} \sum_{k=1}^M d^p \left[\left(P_{0,barycenter}, \{\mu_{i,barycenter}\}_{i=1}^{N-1} \right), \left(\bar{P}_{0,k}, \{\bar{\mu}_{i,k}\}_{i=1}^{N-1} \right) \right]$$



The computation of the Fréchet barycenter is calculated by the Hermann Karcher Flow

Karcher Flow to compute geodesic Mean or Median

Karcher Flow

➤ From Euclidean Case to Manifold Case by exponential map for MEDian

$$m_{n+1} = t \sum_k \frac{x_k - m_n}{\|x_k - m_n\|}$$

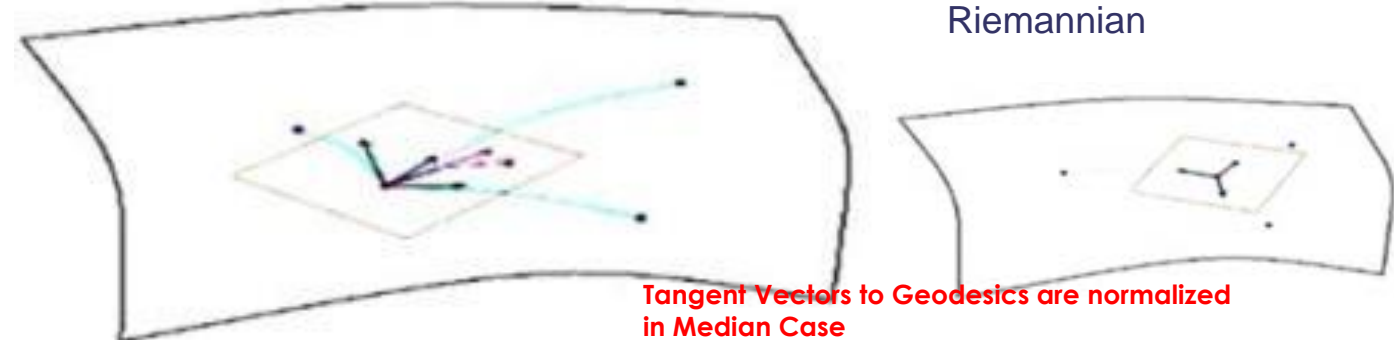
Euclidian

$$m_{n+1} = \exp_{m_n} \left(t \sum_k \frac{\exp_{m_n}^{-1}(x_k)}{\|\exp_{m_n}^{-1}(x_k)\|} \right)$$

Riemannian

Riemannian Center of Mass and
Mollifier Smoothing*

H. KARCHER
Bonn University



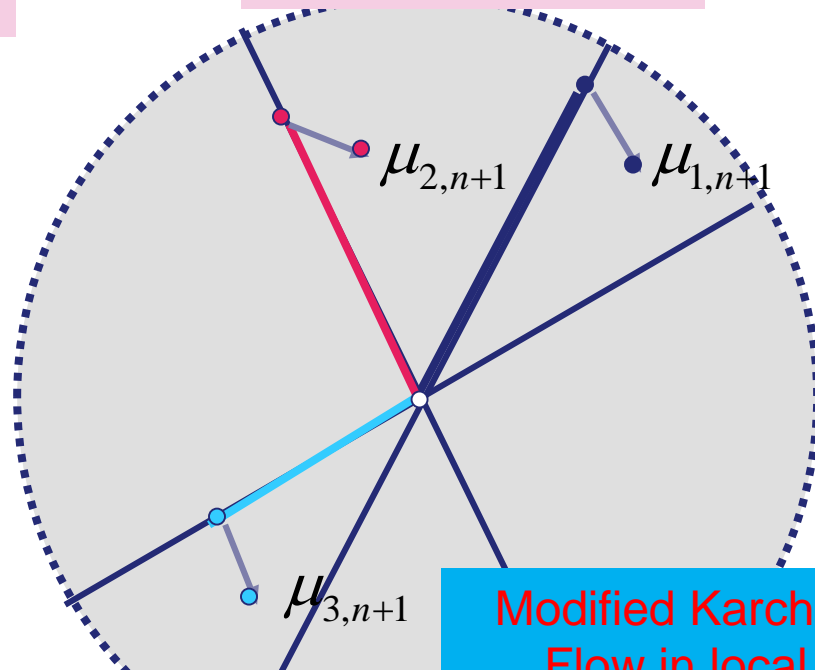
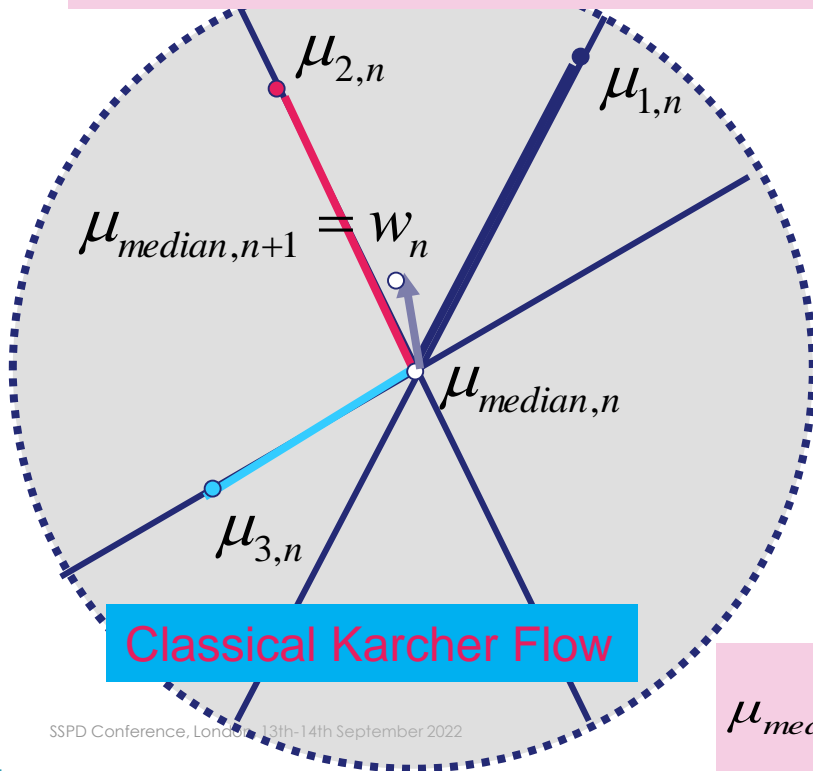
- Starting from an arbitrary point on the Manifold
- Compute the geodesics from this point to the N points and the tangent vectors (normalized)
- Sum of all the Geodesics tangent vectors (normalized)
- Move the point on the manifold in the direction of this sum vector s (via the exponential map)

Hermann KARCHER

Karcher flow for the Fréchet barycenter in the Poincaré unit disk: Computation of reflection / Verblunsky coefficients

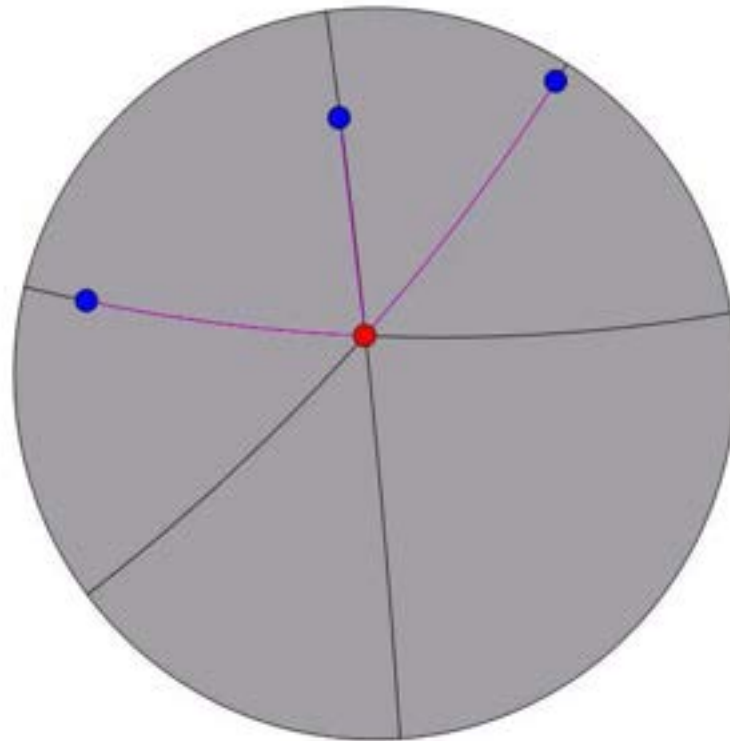
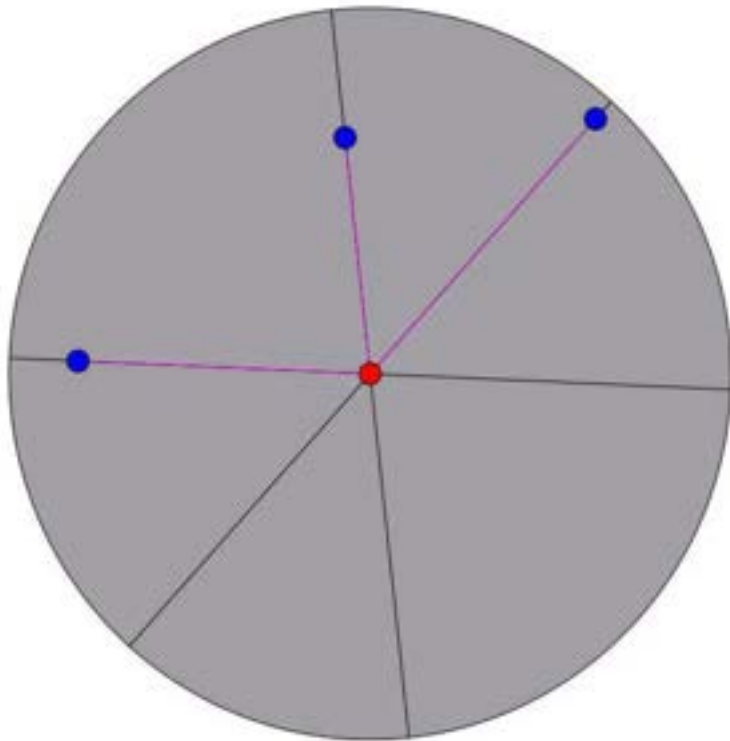
$$w_n = \gamma_n \sum_{\substack{k=1 \\ k \neq l}}^m \frac{\mu_{k,n}}{|\mu_{k,n}|} \quad \text{avec} \quad \{l \mid |\mu_{l,n}| < \varepsilon\}$$

$$\mu_{k,n+1} = \frac{\mu_{k,n} - w_n}{1 - \mu_{k,n} \cdot w_n^*}$$

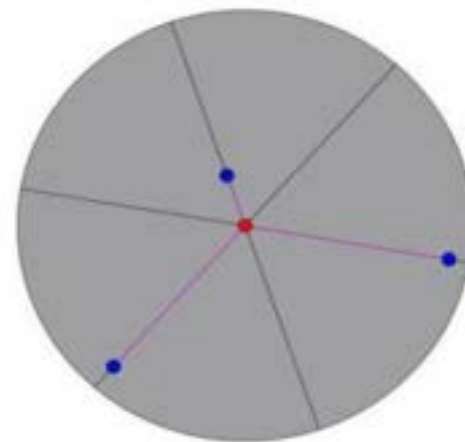
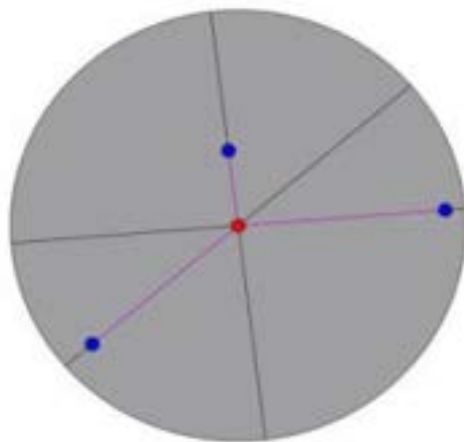
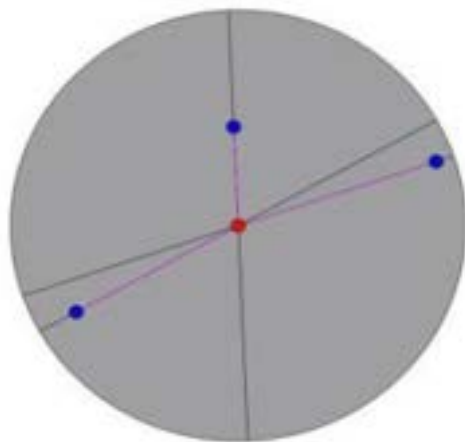
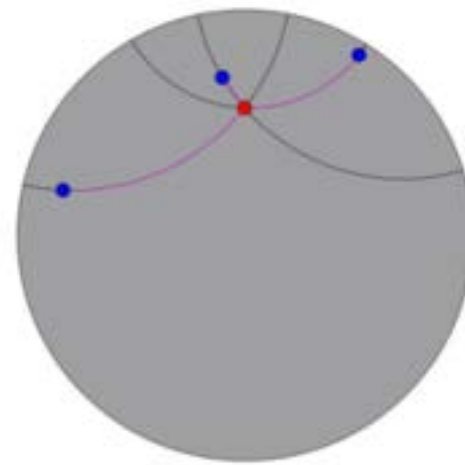
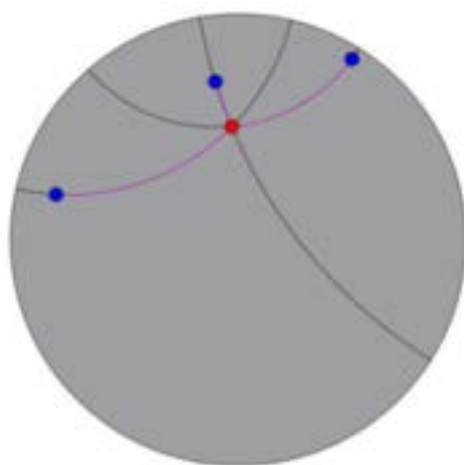
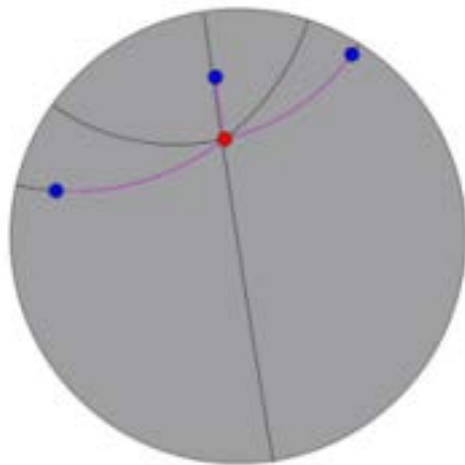


$$\mu_{median,n+1} = \frac{\mu_{median,n} + w_n}{1 + \mu_{median,n} w_n^*}$$

Karcher flow for the Fréchet barycenter in the Poincaré unit disk: Computation of reflection / Verblunsky coefficients

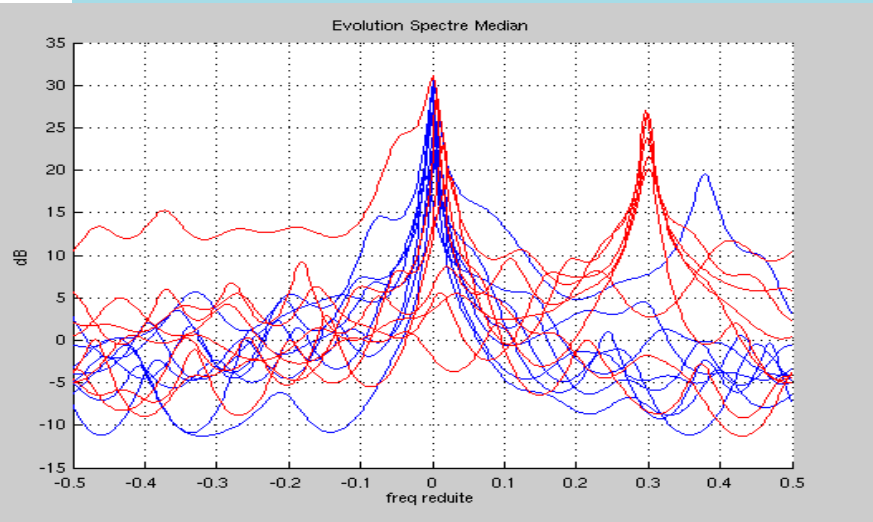


Karcher flow for the Fréchet barycenter in the Poincaré unit disk: Computation of reflection / Verblunsky coefficients

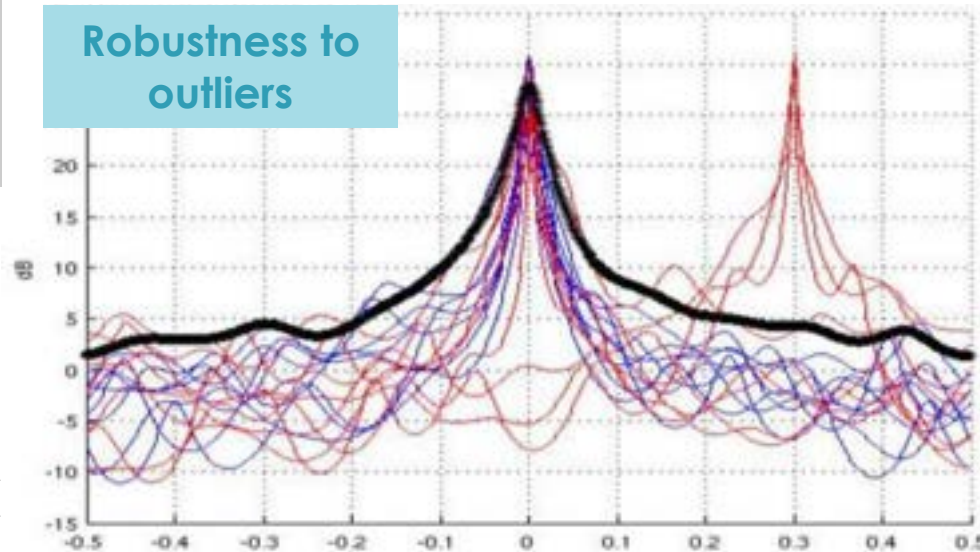


Doppler spectrum "median" by the computation of the Fréchet barycenter on the coefficients of reflection / Verblunsky in the unit disk

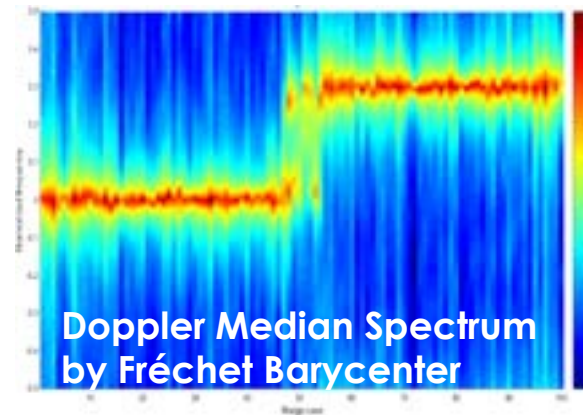
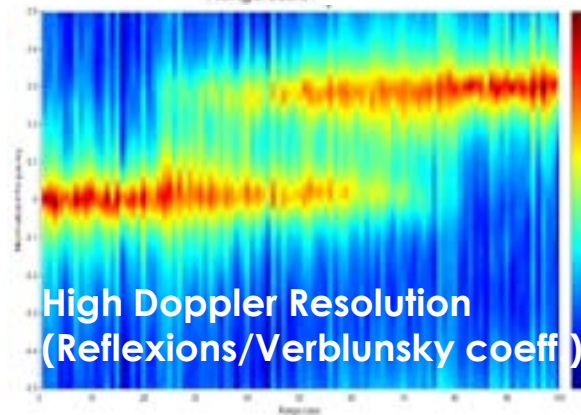
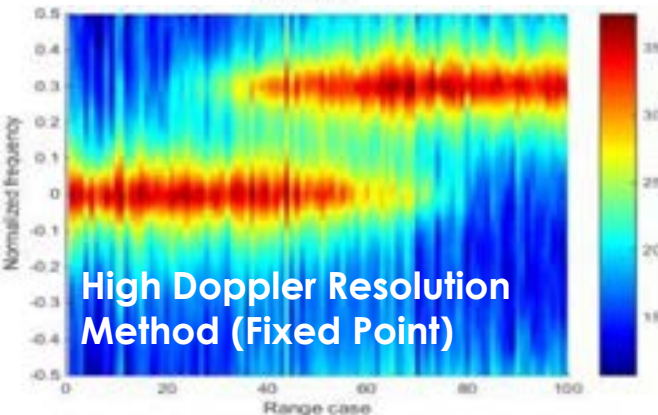
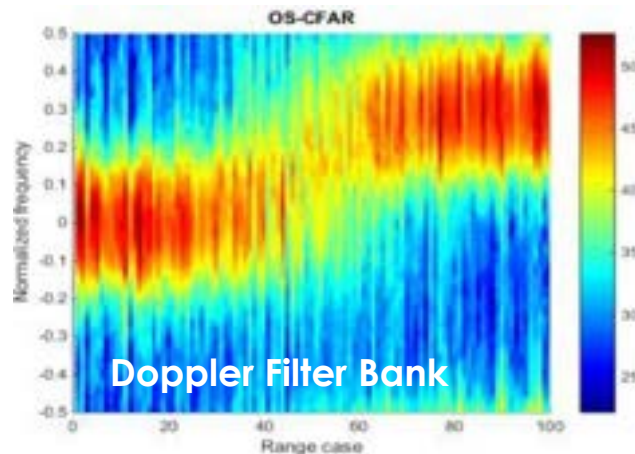
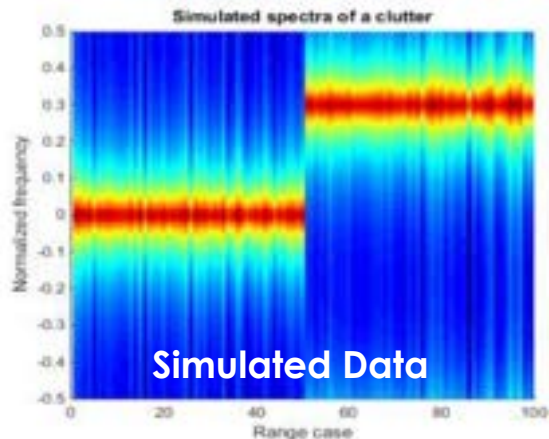
Median Geodesic Barycenter of Doppler Spectrum



Robustness to outliers



Median Doppler Spectrum Estimation: Preserves Discontinuities



Comparison between Doppler Spectrum Mean and Median

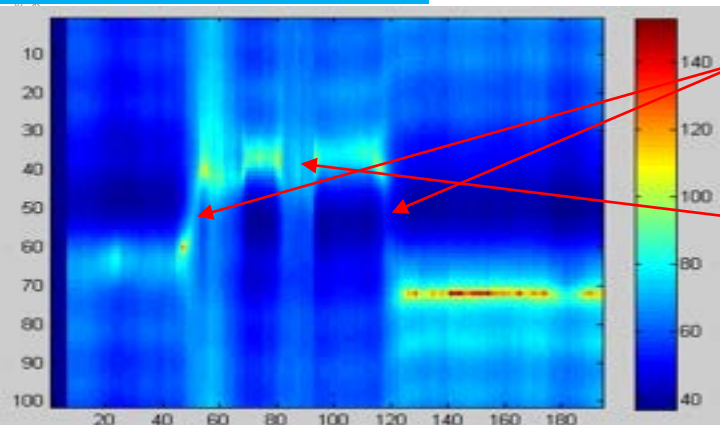
Doppler
axe

Raw Doppler
Spectrum

Distance
axe

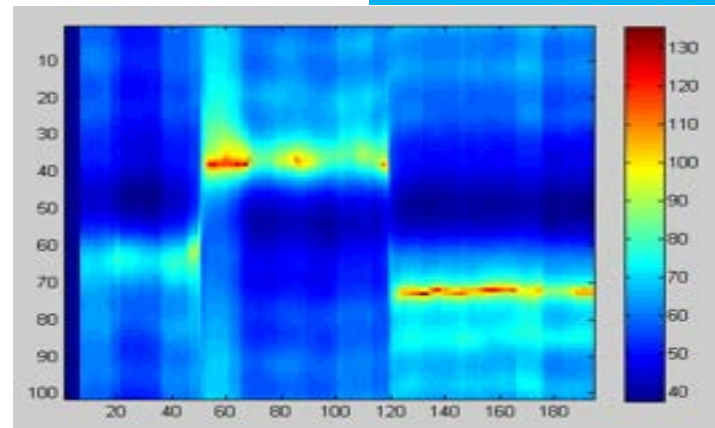
Median Doppler
Spectrum

Mean Doppler
Spectrum



Discontinuities are
not preserved

Perturbation by
Outliers values



Normalized Burg (+ Regularization)

Initialisation

$$P_0 = \frac{1}{N_{cd} N_{rec}} \sum_{i=1}^{N_{cd}} \sum_{k=1}^{N_{rec}} |x_{ik}|^2 \quad \text{For } 1 \leq i \leq N_{cd} \text{ et } 1 \leq n \leq N_{rec} \begin{cases} f_{i,0}(n) = x_{in} \\ b_{i,0}(n) = x_{in} \end{cases}$$

Iteration

$$\text{estimation } \hat{\mu}_m \left[z = \frac{-\frac{2}{N_{cd}} \sum_{i=1}^{N_{cd}} \sum_{n=m+1}^{N_{rec}} \frac{\bar{b}_{i,m-1}(n-1) f_{i,m-1}(n)}{|f_{i,m-1}(n)|^2 + |\bar{b}_{i,m-1}(n-1)|^2} + \gamma \sum_{k=1}^{m-1} a_k^{m-1} a_k^{m-k}}{N_{rec} - m + \gamma \sum_{k=1}^{m-1} |a_k^{m-1}|^2} \right] \quad B_1: x \rightarrow \frac{1-x^2}{x} \left(\frac{\log(1-x) - \log(1+x)}{2x} \right) + \frac{1}{1-x^2} \quad \hat{\mu}_m = B_1^{-1}(|z|) \frac{z}{|z|}$$

$$P_m = (1 - |\hat{\mu}_p|^2) P_{m-1} \begin{pmatrix} a_1^{(m)} \\ \vdots \\ a_{m-1}^{(m)} \end{pmatrix} = \begin{pmatrix} a_1^{(m-1)} \\ \vdots \\ a_{m-1}^{(m-1)} \end{pmatrix} + \hat{\mu}_m \begin{pmatrix} a_1^{(m-1)} \\ \vdots \\ a_{m-1}^{(m-1)} \end{pmatrix} \quad a_m^{(m)} = \hat{\mu}_m \quad \text{For } 1 \leq i \leq N_{cd} \text{ et } m+1 \leq n \leq N_{rec} \begin{cases} f_{i,m}(n) = f_{i,m-1}(n) + \hat{\mu}_m b_{i,m-1}(n-1) \\ b_{i,m}(n) = b_{i,m-1}(n-1) + \hat{\mu}_m f_{i,m-1}(n) \end{cases}$$

Burg Reflection Coefficients Median

Computation of Median

$$\hat{\mu}_{i, median} = \underset{\mu}{\operatorname{argmin}} \left(\sum_{k=1}^{N_{cd}} d(\mu, \mu_{i,k}) \right)$$

$$d(\mu_1, \mu_2) = \frac{1}{2} \log \left(\frac{1 + \left| \frac{\mu_1 - \mu_2}{1 - \bar{\mu}_1 \mu_2} \right|}{1 - \left| \frac{\mu_1 - \mu_2}{1 - \bar{\mu}_1 \mu_2} \right|} \right)$$

Iteration

$$\varphi_w(z) = \frac{z+w}{1+z\bar{w}}$$

$$w_n = \sum_{k \in B_0} \frac{\varphi_{x_n}(z_k)}{|\varphi_{x_n}(z_k)|} \quad B_0 = \{k \in [1, N], |\varphi_{x_n}(z_k)| < \eta\}$$

$$y_{n+1} = \tanh(|t_n w_n|) \frac{t_n w_n}{|t_n w_n|}$$

$$x_{n+1} = \varphi_{-x_n}(y_{n+1})$$

Critère arrêt

$$|w_n| < \epsilon$$

Deetector based on median covariance matrix

2 types de détecteurs ont été testés:

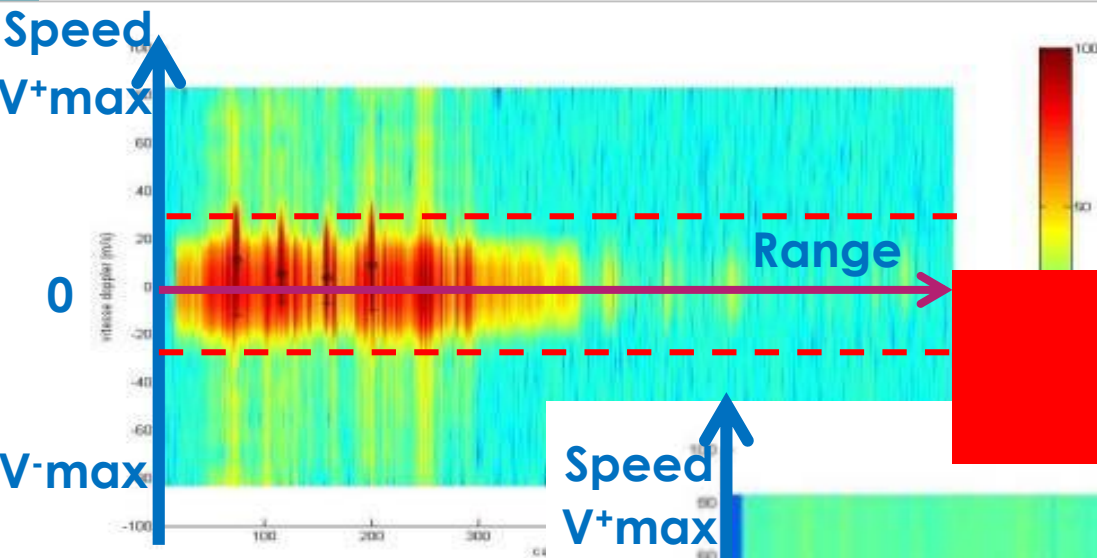
- **Normalized GLR Detector** (classical GLR Detector, normalized to be independant of amplitude): test for Different Doppler values

$$GLRT(z) = \max_{\theta \in [-0.5; 0.5[} \frac{|p(\theta) + \hat{\Sigma}^{-1}z|^2}{(z + \hat{\Sigma}^{-1}z)(p(\theta) + \hat{\Sigma}^{-1}p(\theta))} \quad p(\theta) = (1, e^{2i\pi\theta}, \dots, e^{2i\pi(d-1)\theta})^T$$

- **Geometric Detector:** Information Geometry Fisher Rao Distance between reflection Coefficient between cell under test and Median values of Ambiance

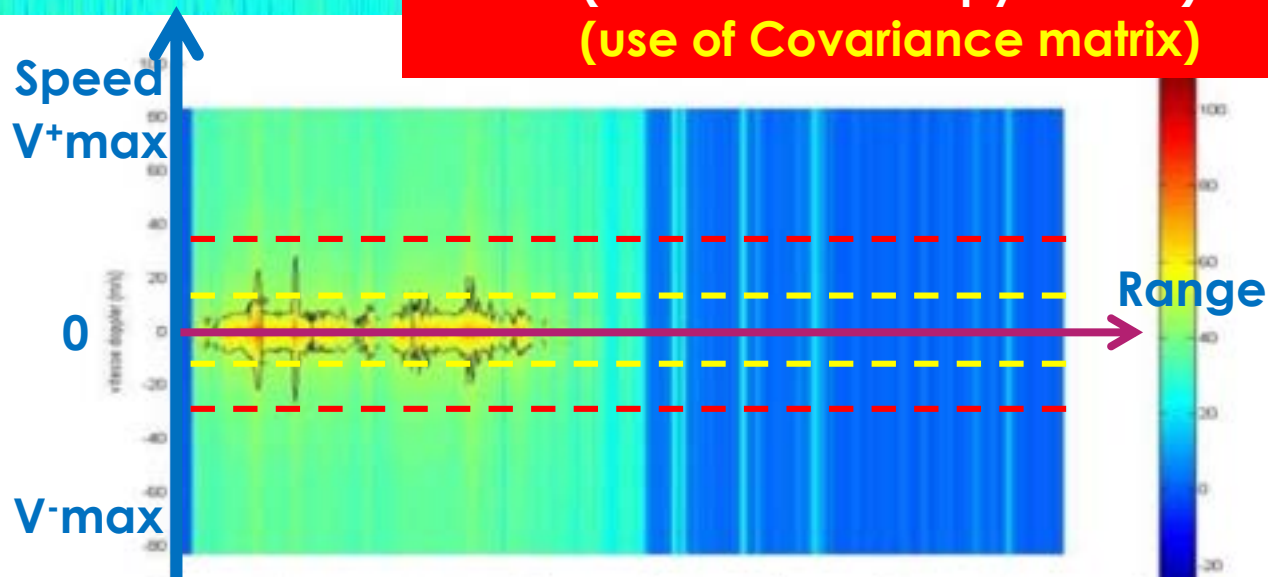
$$AR(z) = \sum_{k=1}^M (M - k) d(\hat{\mu}_k(z), \hat{\mu}_{k,amb})$$

MATRIX Geometric CFAR

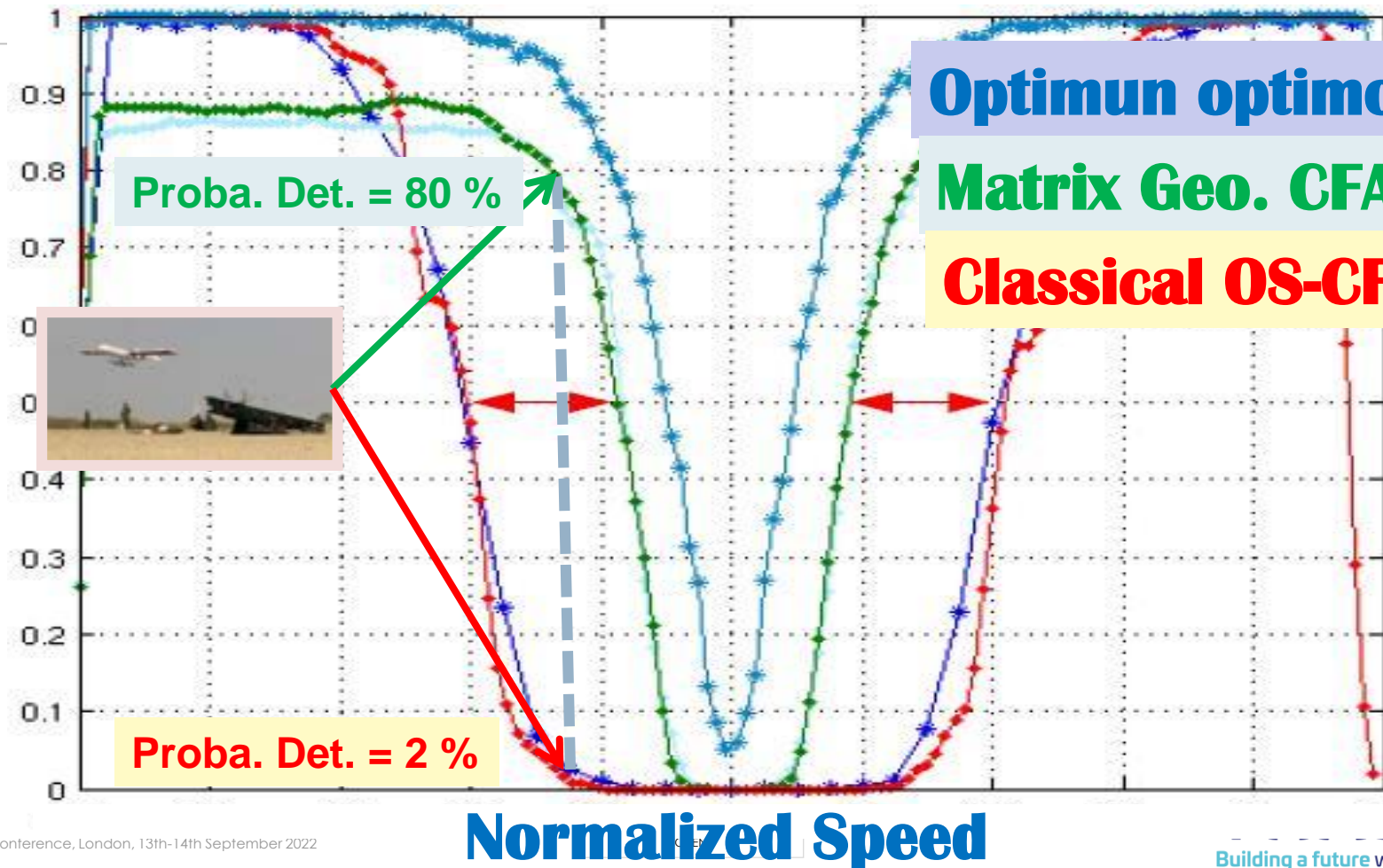


High Doppler Resolution
(Maximum Entropy Model)
(use of Covariance matrix)

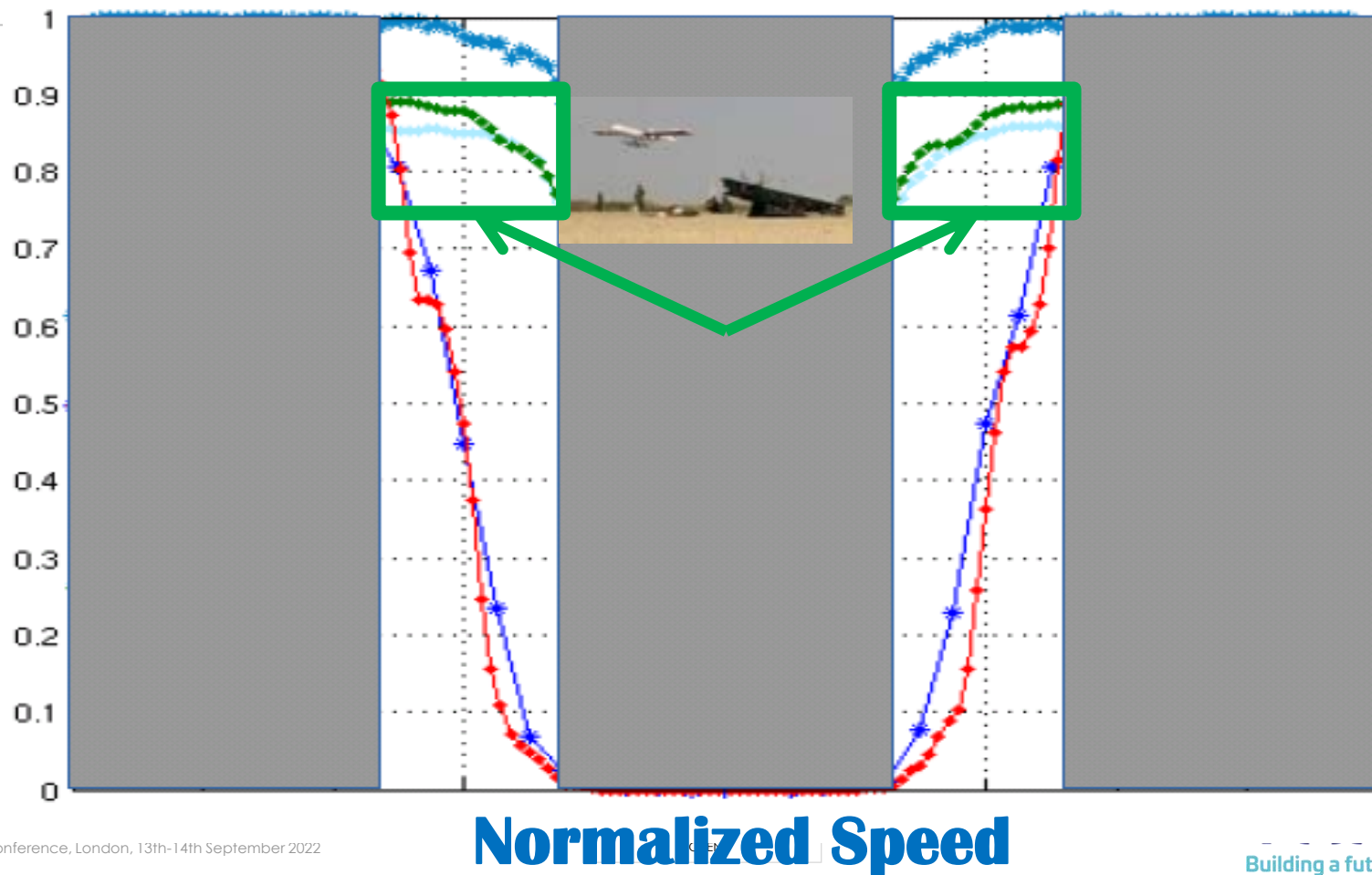
OS-HDR-CFAR



Probability of Detection

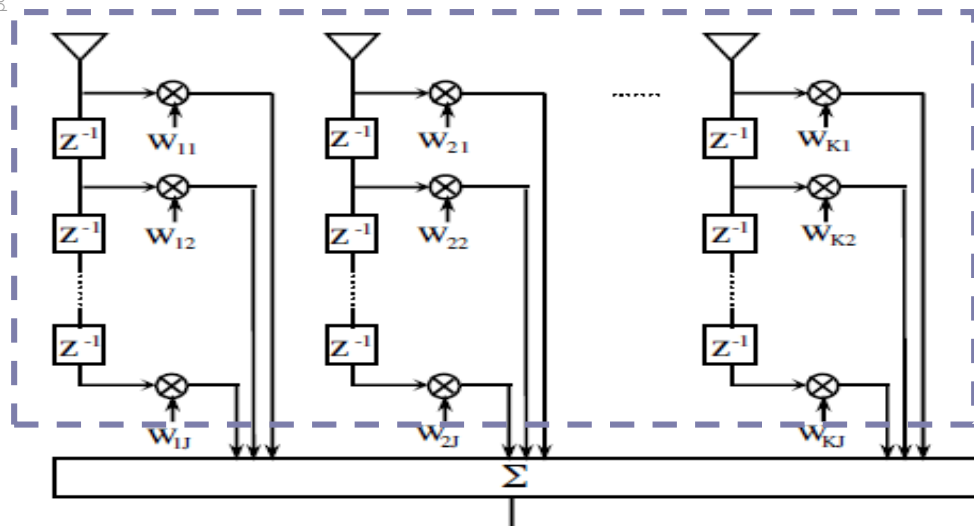


Probability of Detection



Extension to the spatio-temporal measurement of the electromagnetic wave: Toeplitz-Bloc-Toeplitz matrix structure

We consider a vector of spatial and temporal measures



Digital Space-Time Measure
(Space-Time covariance matrix)

$z_{i,j} :$ $\begin{cases} i: \text{Time index} \\ j: \text{Spatial index} \end{cases}$

$$R = E[ZZ^+]$$

$$Z = \begin{bmatrix} z_{1,1} \\ \vdots \\ z_{N,1} \\ \vdots \\ z_{1,M} \\ \vdots \\ z_{N,M} \end{bmatrix}$$

$$R_{p,n+1} = \begin{bmatrix} R_0 & R_1 & \cdots & R_n \\ R_1^+ & R_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & R_1 \\ R_n^+ & \cdots & R_1^+ & R_0 \end{bmatrix} = \begin{bmatrix} R_{p,n} & \tilde{R}_n \\ \tilde{R}_n^+ & R_0 \end{bmatrix}$$

$$\tilde{R}_n = V \begin{bmatrix} R_1 \\ \vdots \\ R_n \end{bmatrix}^* \text{ with } V = \begin{bmatrix} 0 & \cdots & 0 & J_p \\ \vdots & \ddots & \ddots & 0 \\ 0 & J_p & \ddots & \vdots \\ J_p & 0 & \cdots & 0 \end{bmatrix}$$

Extension to Toeplitz Block Toeplitz Hermitian PD Matrices

Considering Toeplitz-Block-Toeplitz Matrices :

$$R_{p,n+1} = \begin{bmatrix} R_0 & R_1 & \cdots & R_n \\ R_1^+ & R_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & R_1 \\ R_n^+ & \cdots & R_1^+ & R_0 \end{bmatrix} = \begin{bmatrix} R_{p,n} & \tilde{R}_n \\ \tilde{R}_n^+ & R_0 \end{bmatrix}$$
$$\tilde{R}_n = V \begin{bmatrix} R_1 \\ \vdots \\ R_n \end{bmatrix}^* \quad \text{with} \quad V = \begin{bmatrix} 0 & \cdots & 0 & J_p \\ \vdots & \ddots & \ddots & 0 \\ 0 & J_p & \ddots & \vdots \\ J_p & 0 & \cdots & 0 \end{bmatrix}$$

Efficient Inversion of Toeplitz-Block Toeplitz Matrix

MATI WAX AND THOMAS KAILATH, FELLOW, IEEE

Extension of the Trench / Verblunsky theorem to the case of Hermitian Toeplitz-Bloc-Toeplitz Positive Definite Matrices

Positive Definite Hermitian Toeplitz-Bloc-Toeplitz matrices can be parameterized by Matrix Verblunsky coefficients:

$$R_{p,n+1}^{-1} = \begin{bmatrix} \alpha_n & \alpha_n \cdot \widehat{A}_n^+ \\ \alpha_n \cdot \widehat{A}_n & R_{p,n}^{-1} + \alpha_n \cdot \widehat{A}_n \cdot \widehat{A}_n^+ \end{bmatrix} \quad \text{with} \quad \alpha_n^{-1} = [1 - A_n^n A_n^{n+}] \alpha_{n-1}^{-1}, \quad \alpha_0^{-1} = R_0$$

$$R_{p,n+1} = \begin{bmatrix} \alpha_n^{-1} + \widehat{A}_n^+ \cdot R_{p,n} \cdot \widehat{A}_n & -\widehat{A}_n^+ \cdot R_{p,n} \\ -R_{p,n} \cdot \widehat{A}_n & R_{p,n} \end{bmatrix} \quad \text{and} \quad \widehat{A}_n = \begin{bmatrix} A_1^1 \\ \vdots \\ A_n^n \end{bmatrix} = \begin{bmatrix} \widehat{A}_{n-1} \\ 0_p \end{bmatrix} + A_n^n \cdot \begin{bmatrix} J_p A_{n-1}^{n-1*} J_p \\ \vdots \\ J_p A_1^{n-1*} J_p \\ I_p \end{bmatrix}$$

Extension of the Trench / Verblunsky theorem to the matrix case: Existence of a diffeomorphism φ :

$$\varphi : TBTHPD_{n \times n} \rightarrow THPD_n \times SD^{n-1}$$

$$R \mapsto (R_0, A_1^1, \dots, A_{n-1}^{n-1})$$

$$\text{with } SD = \{Z \in Herm(n) / ZZ^+ < I_n\}$$

Coding Space-Time Data in Siegel Poly-Disk

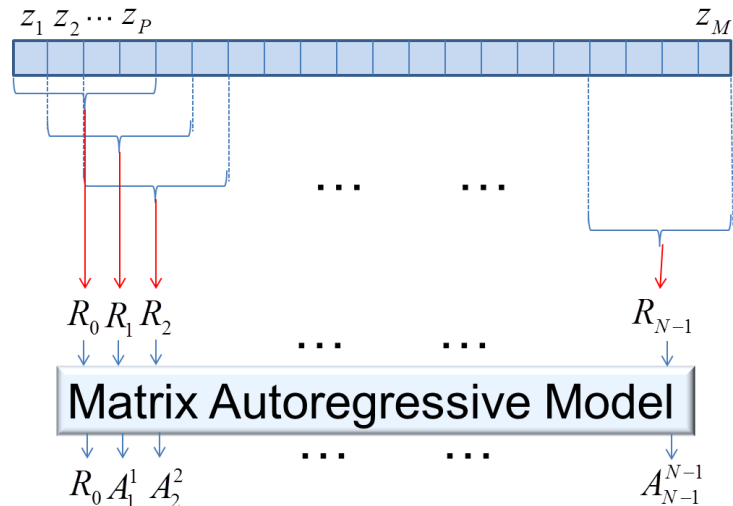
$$\{z_1, \dots, z_p, \dots, z_M\} \xRightarrow{\text{split}} \left\{ \{z_1, \dots, z_p\}, \{z_{p+1}, \dots, z_{2p}\}, \dots, \{z_{M-p+1}, \dots, z_M\} \right\}$$

Time series $\{R_0, \dots, R_{N-1}\}$ as TBTHPD matrix

$$R_{N,p} = \begin{bmatrix} R_0 & R_1 & \dots & R_{N-1} \\ R_1^+ & R_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & R_1 \\ R_{N-1}^+ & \dots & R_1^+ & R_0 \end{bmatrix}$$

Matrix Verblunsky/Trench Theorem

$$\{R_0, A_1^1, \dots, A_{N-1}^{N-1}\} \text{ with } A_k^k \in SD = \{Z / ZZ^+ < I\}$$



Metric of Information Geometry: Hessian of Potential

- Entropy defines a (Kählerian) potential whose Hessian provides a Riemannian metric for Verblunsky matrix parameters:

$$\begin{aligned}\tilde{\Phi}(R_{p,n}) &= -\log(\det R_{p,n}) + cste = -Tr(\log R_{p,n})\mu + cste \\ \Rightarrow g_{ij} &= Hess[\phi(R_{p,n})]\end{aligned}$$

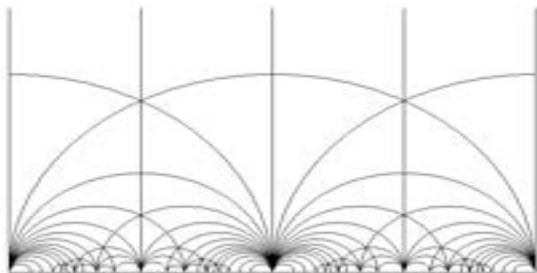
- The "Toeplitz-Bloc-Toeplitz" structure allows to express the entropy only by the Verblunsky parameters about spatial information and of a matrix R_0 incorporating the Doppler information (time information) as previously studied:

$$\tilde{\Phi}(R_{p,n}) = \sum_{k=1}^{n-1} (n-k) \cdot \log \det [I_n - A_k^k A_k^{k+}] + n \cdot \log [\pi \cdot e \cdot \det R_0]$$

- The Hessian of Entropy will provide the following metric:

$$ds^2 = n \cdot Tr \left[(R_0^{-1} dR_0)^2 \right] + \sum_{k=1}^{n-1} (n-k) Tr \left[(I_n - A_k^k A_k^{k+})^{-1} dA_k^k (I_n - A_k^{k+} A_k^k)^{-1} dA_k^{k+} \right]$$

Extension of homogeneous bounded symmetric domains: Siegel Upper half-space and Siegel disk



Poincaré Upper half Plane

$$ds^2 = \frac{dx^2 + dy^2}{y^2} = \frac{|dz|^2}{y^2}$$

$$ds^2 = y^{-1} dz y^{-1} dz^*$$

with $z = x + iy$ and $y > 0$

Siegel Upper Half Space

$$ds^2 = \text{Tr}(Y^{-1} dZY^{-1} d\bar{Z})$$

with $Z = X + iY$

⁷⁴ $X \in \text{Herm}(n, \mathbb{C})$ and $Y \in \text{HPD}(n, \mathbb{C})$ □

$$w = \frac{z - i}{z + i}$$



Poincaré Unit Disk

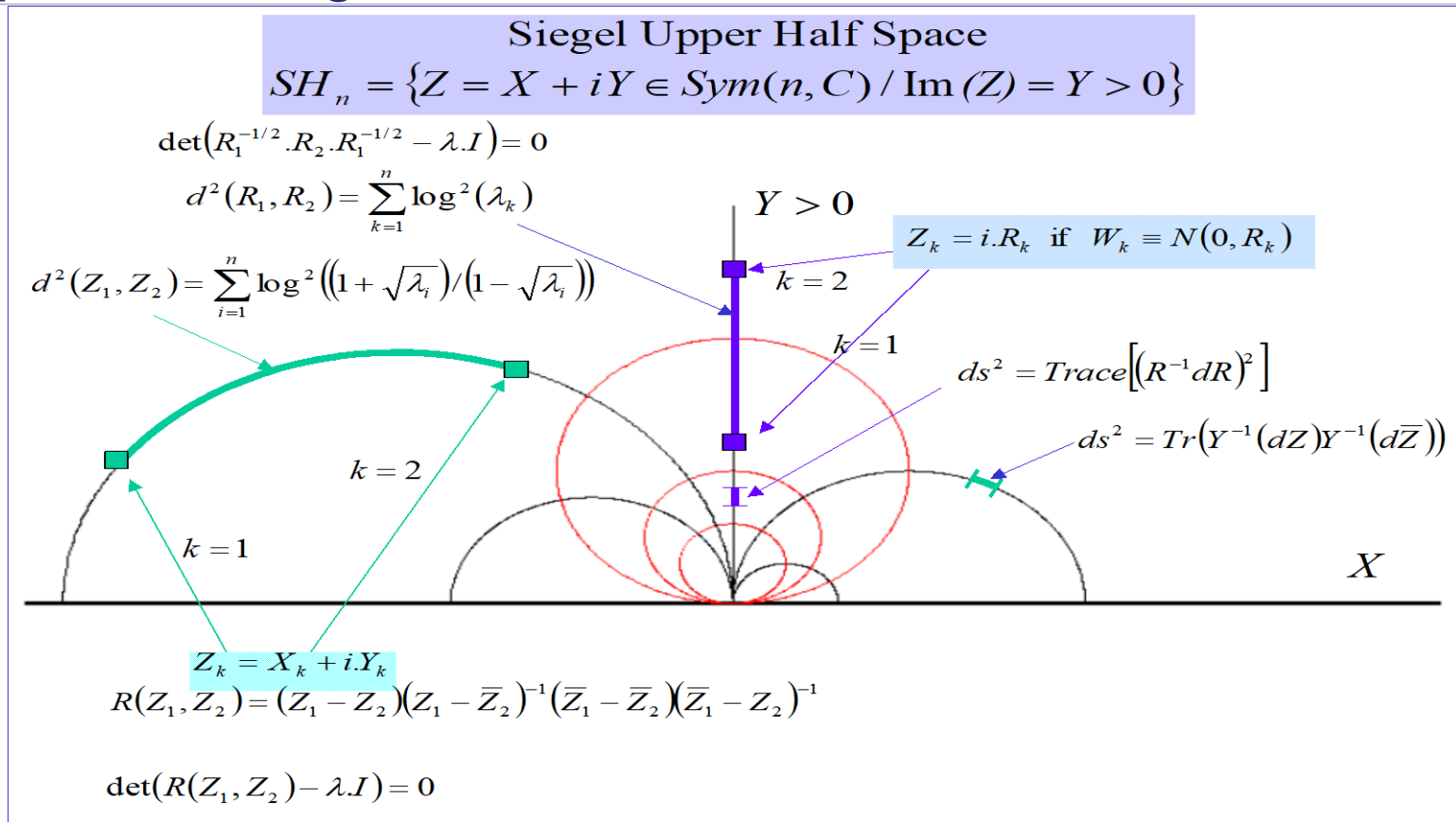
$$ds^2 = \frac{|dw|^2}{(1 - |w|^2)^2}$$

$$ds^2 = (1 - ww^*)^{-1} dw (1 - ww^*)^{-1} dw^*$$

Siegel Unit Disk

$$ds^2 = \text{Tr}[(I - WW^+)^{-1} dW (I - W^+W)^{-1} dW^+]$$

Extension of homogeneous Bounded symmetric domains: Siegel Upper half-space and Siegel disk



Siegel Disc Automorphism

➤ Automorphism of Siegel Disc SD_n given by :

$$\Sigma = \Phi_{Z_0}(Z) = (I - Z_0 Z_0^+)^{-1/2} (Z - Z_0) (I - Z_0^+ Z)^{-1} (I - Z_0^+ Z_0)^{1/2}$$

➤ All automorphisms given by :

$$\forall \Psi \in \text{Aut}(SD_n), \exists U \in U(n, C) / \Psi(Z) = U \Phi_{Z_0}(Z) U^t$$

➤ Distance given by:

$$\forall Z, W \in SD_n, d(Z, W) = \frac{1}{2} \log \left(\frac{1 + \|\Phi_Z(W)\|}{1 - \|\Phi_Z(W)\|} \right)$$

➤ Inverse automorphism given by :

$$G = (I - Z_0 Z_0^+)^{1/2} \Sigma (I - Z_0^+ Z_0)^{-1/2} = (Z - Z_0) (I - Z_0^+ Z)^{-1}$$
$$\Rightarrow \begin{cases} Z = \Phi_{Z_0}^{-1}(\Sigma) = (G Z_0^+ + I)^{-1} (G + Z_0^+) \\ \text{with } G = (I - Z_0 Z_0^+)^{1/2} \Sigma (I - Z_0^+ Z_0)^{-1/2} \end{cases}$$

OPEN

Iterated Computation of Median in Siegel disk

Initialisation : $W_{median,0} = 0$ and $\{W_{1,0}, \dots, W_{m,0}\} = \{W_1, \dots, W_m\}$

Iterate on n until $\|G_n\|_F < \varepsilon$

$$W_{k,n} = U_{k,n} e^{iA_{k,n}} e^{S_{k,n}} \Rightarrow H_{k,n} = U_{k,n} e^{iA_{k,n}} = W_{k,n} e^{-S_{k,n}} = e^{-\frac{S_{k,n}}{2}} W_{k,n} e^{-\frac{S_{k,n}}{2}} \text{ with:}$$

$$S_{k,n} = 1/2 \cdot \log \left(P_{k,n}^{1/2} \left(P_{k,n}^{-1/2} P_{k,n}^* P_{k,n}^{-1/2} \right)^{1/2} P_{k,n}^{1/2} \right) \text{ with } P_{k,n} = W_{k,n}^+ W_{k,n}$$

$$G_n = \gamma_n \sum_{\substack{k=1 \\ k \neq l}}^m H_{k,n} \text{ with } \{l / \|H_{k,n}\|_F < \varepsilon\}$$

For $k = 1, \dots, m$ then $W_{k,n+1} = \Phi_{G_n}(W_{k,n})$

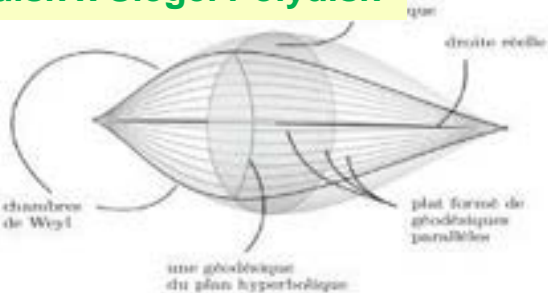
$$W_{k,n+1} = (I - G_n G_n^+)^{-1/2} (W_{k,n} - G_n) (I - G_n^+ W_{k,n})^{-1} (I - G_n^+ G_n)^{1/2}$$

$$W_{median,n+1} = (G G_n^+ + I)^{-1} (G + G_n) \text{ with } G = (I - G_n G_n^+)^{1/2} W_{median,n} (I - G_n^+ G_n)^{-1/2}$$

Coding of Digital Electromagnetic Wave States in $R \times S^1 \times D^{m-1} \times SD^{n-1}$

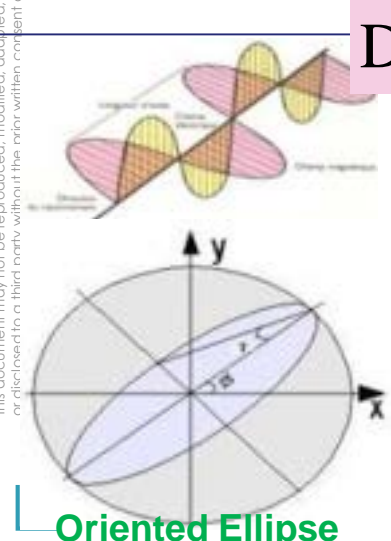
$$\begin{aligned} (R_0, A_1^1, \dots, A_{n-1}^{n-1}) &\in THPD_m \times SD^{n-1} \\ SD &= \{Z / ZZ^+ < I_m\} \\ R_0 &\rightarrow (\log(P_0), \mu_1, \dots, \mu_{m-1}) \in R \times D^{m-1} \\ D &= \{z / zz^* < 1\} \end{aligned}$$

Space-Time Information:
 $R \times$ Poincaré Polydisk \times Siegel Polydisk

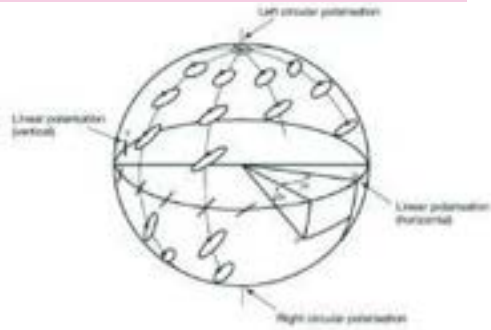


Spatio-Doppler Coding $\in R \times D^{m-1} \times SD^{n-1}$

Digital Electromagnetic Wave Code $\in R \times S^1 \times D^{m-1} \times SD^{n-1}$



$$\vec{S} = \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} s_0 \\ s_0 \cos 2\tau \cos 2\phi \\ s_0 \cos 2\tau \sin 2\phi \\ s_0 \sin 2\tau \end{bmatrix} \text{ avec } \begin{cases} \phi = \frac{1}{2} \arctan\left(\frac{s_2}{s_1}\right) \\ \tau = \frac{1}{2} \arctan\left(\frac{s_3}{\sqrt{s_1^2 + s_2^2}}\right) \end{cases}$$



Polarimetric Coding $\in R \times S^1$

Polarimetric Information:
 $R_+^* \times S^1$

This document may not be reproduced, modified, adapted, or disseminated in a third party without the prior written consent of the author.

Poincaré geometric modeling of polarization states

Polarization state representation by a point in a complex plane

➤ Polarisation Ellipse

$$E_M = \begin{bmatrix} E_{M,x}(t) \\ E_{M,y}(t) \end{bmatrix} = \begin{bmatrix} A_x(t) \cdot e^{i\varphi_x(t)} \\ A_y(t) \cdot e^{i\varphi_y(t)} \end{bmatrix} = p e^{i\varphi} \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix} \begin{bmatrix} \cos \tau \\ i \sin \tau \end{bmatrix}$$

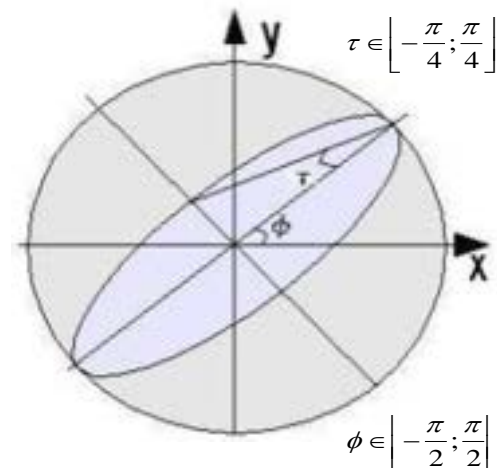
- τ : angle of ellipticity characterizing the elongation (the shape)
- ϕ : characterizes the inclination of the ellipse

➤ Jones Matrix:

$$J = E(E_M \cdot E_M^+) = \begin{bmatrix} E(|E_{M,x}|^2) & E(E_{M,y} E_{M,x}^*) \\ E(E_{M,x} E_{M,y}^*) & E(|E_{M,y}|^2) \end{bmatrix} = \begin{bmatrix} J_{11} & J_{21} \\ J_{12} & J_{22} \end{bmatrix}$$

➤ Stokes Vectors:

$$\vec{S} = \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} J_{11} + J_{22} \\ J_{11} - J_{22} \\ 2\text{Re}(J_{21}) \\ -2\text{Im}(J_{12}^{\text{OPEN}}) \end{bmatrix} = \begin{bmatrix} s_0 \\ s_0 \cos 2\tau \cos 2\phi \\ s_0 \cos 2\tau \sin 2\phi \\ s_0 \sin 2\tau \end{bmatrix}$$



$$\phi = \frac{1}{2} \arctan \left(\frac{s_2}{s_1} \right)$$

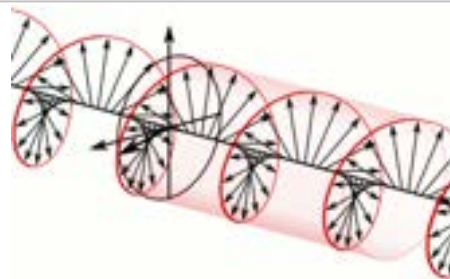
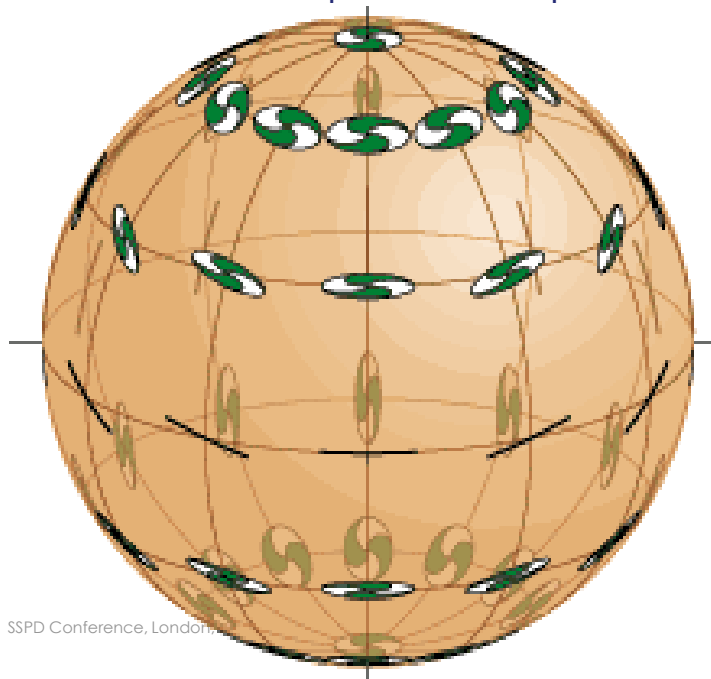
$$\tau = \frac{1}{2} \arctan \left(\frac{s_3}{\sqrt{s_1^2 + s_2^2}} \right)$$

$$s_0^2 \geq s_1^2 + s_2^2 + s_3^2$$

Poincaré geometric modeling of polarization states

Spherical coordinates

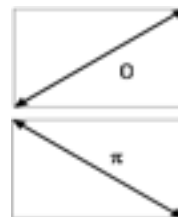
- On the sphere: $(2\phi, 2\tau)$
- Northern hemisphere: right polarisation
- Southern hemisphere: left polarisation



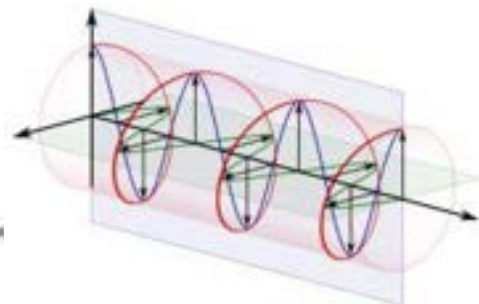
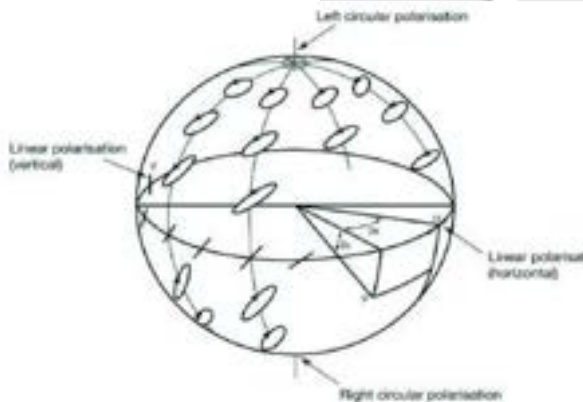
Left Circular Polarization



Right Circular Polarization

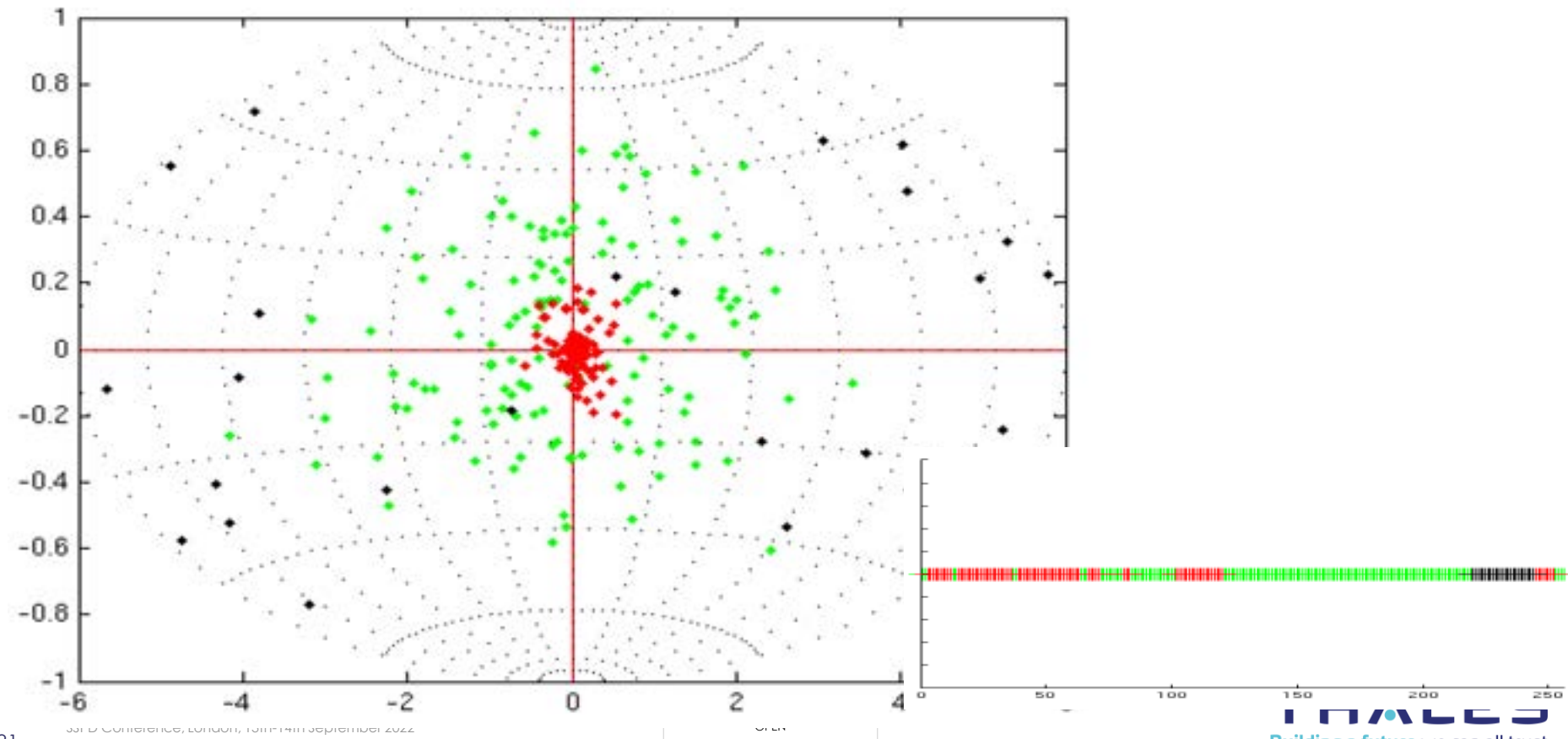


Left circular polarisation



Segmentation/classification of Polatrimetric states

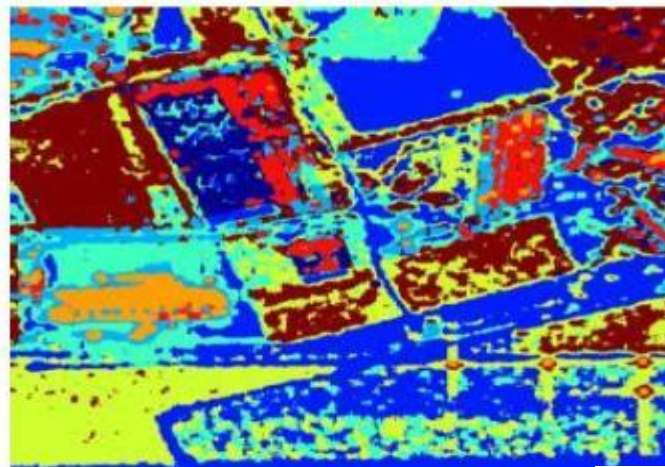
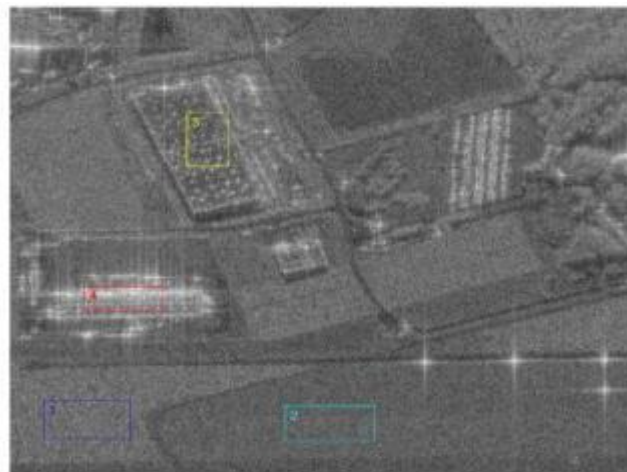
This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of THALES - © 2021 THALES. All rights reserved.



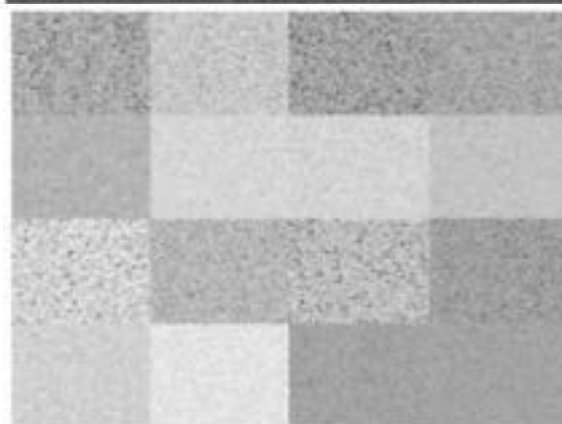
33rd Conference, London, 13th-14th September 2024

CHES

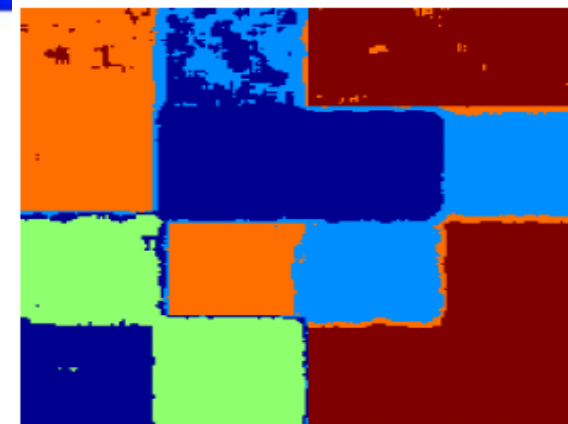
Pierre Formont (ONERA PhD): Segmentation of Polarimetric Data



Segmentation based
on distance from
Information
Geometry



| | | | |
|------------------------------------|------------------------------------|------------------------------------|------------------------------------|
| $\sim \mathcal{K}(M_1, \lambda_1)$ | $\sim \mathcal{K}(M_1, \lambda_2)$ | $\sim \mathcal{K}(M_2, \lambda_1)$ | $\sim \mathcal{K}(M_2, \lambda_2)$ |
| $\sim \mathcal{K}(M_1, \lambda_3)$ | $\sim \mathcal{K}(M_1, \lambda_4)$ | $\sim \mathcal{K}(M_2, \lambda_3)$ | $\sim \mathcal{K}(M_2, \lambda_4)$ |
| $\sim \mathcal{K}(M_3, \lambda_1)$ | $\sim \mathcal{K}(M_3, \lambda_2)$ | $\sim \mathcal{K}(M_4, \lambda_1)$ | $\sim \mathcal{K}(M_4, \lambda_2)$ |
| $\sim \mathcal{K}(M_3, \lambda_3)$ | $\sim \mathcal{K}(M_3, \lambda_4)$ | $\sim \mathcal{K}(M_4, \lambda_3)$ | $\sim \mathcal{K}(M_4, \lambda_4)$ |



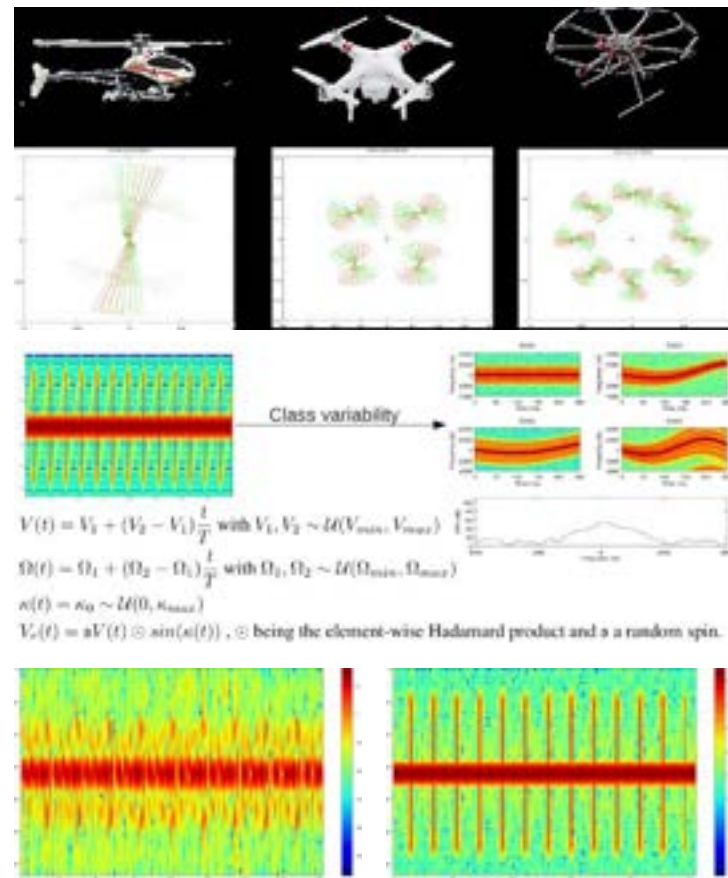
SCM - GM

Advanced Sensor Processing based on Koszul Information Geometry

- Geometric Matrix CFAR/STAP for (very) slow targets detection in clutter
- Complex-Valued CNN & Covariance-Matrix-Valued HPDNet for Micro-Doppler ATDR
- Tracker parameters tuning by Deep Learning for tracking hyper-maneuvering targets
- Multi-Agent Reinforcement Learning for Sensor Resources Management for tracking hyper-maneuvering targets
- Multi-Sensors Collaborative Tracking by Distributed Auctions for tracking in saturating scenario (swarm, fleet of targets, ...)

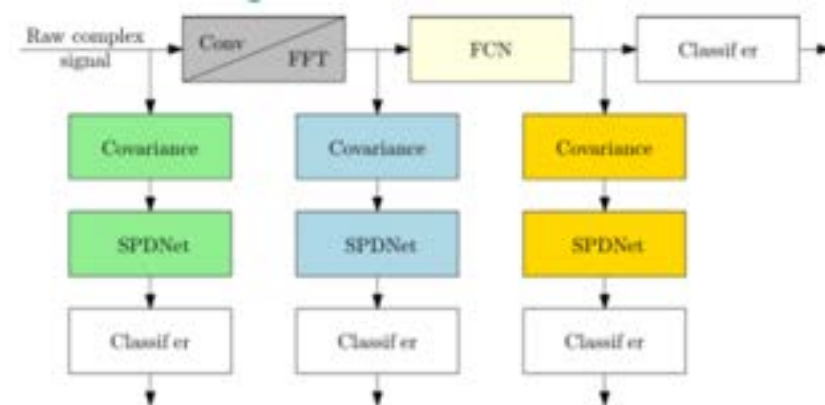
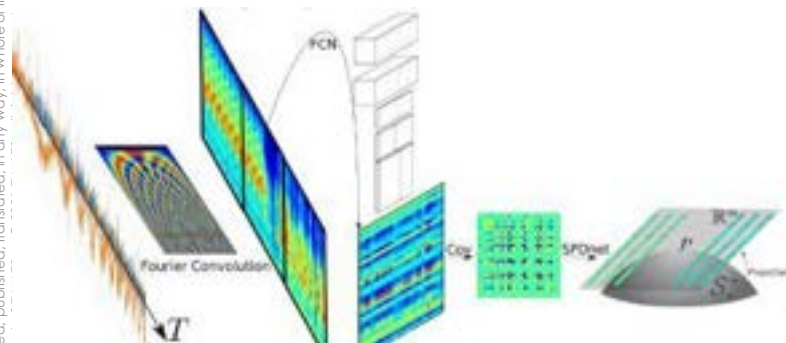
Drone Recognition by Radar Micro-Doppler Signature

- **Drone Radar Micro-Doppler Deep Learning**
 - Use of Radar simulator to learn on Hybrid data (simulated and real data)
 - Micro-Doppler time/frequency spectrum
- **Complex-valued CNN**
 - Fourier transform is a convolution by the Fourier atoms. We can learn a Fourier-like complex filter bank.
- **HPD neural networks**
 - Covariance matrix has HPD (Hermitian Positive Definite) structure
 - Statistical analysis of manifold-valued data : Information Geometry

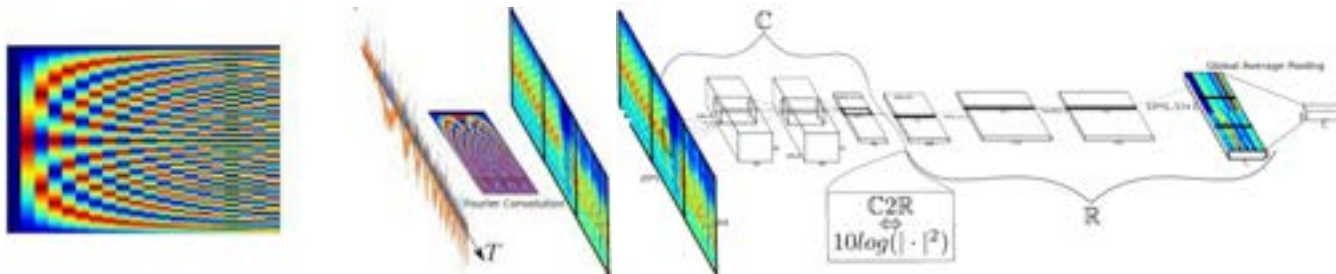


Fully CNN, SPDNet/HPDNet Architecture and Complex Processing

Adaptation for SPD/HPD matrix

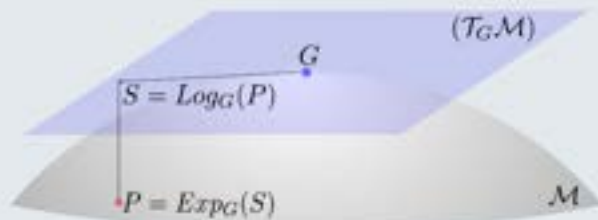


Adaptation for Complex convolution & Fully Convolution Network (time axis)



SPDNET & HPNET

Tools on SPD manifold: Tangent space

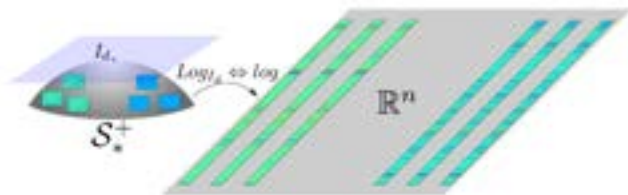


Exp and Log manifold mappings.

Machine Learning consequence

Possible to implement learning methods in a Euclidean setting on the tangent space of some given reference point.

→ Code Riemannian method using projection and the Euclidean method. → Choice of reference projection point.



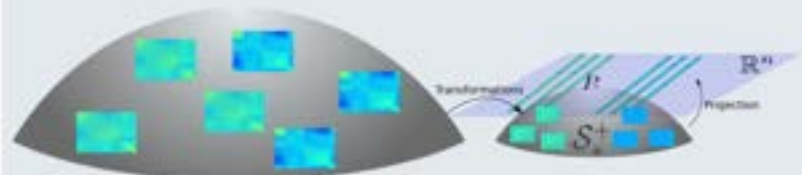
Tools on SPD manifold: Manifold mappings

Mapping operator from manifold to tangent space, and its inverse:

$$\forall S \in \mathcal{T}_G, \text{Exp}_G(S) = G^{\frac{1}{2}} \exp(G^{-\frac{1}{2}} S G^{-\frac{1}{2}}) G^{\frac{1}{2}} \in \mathcal{M}$$

$$\forall P \in \mathcal{M}, \text{Log}_G(P) = G^{\frac{1}{2}} \log(G^{-\frac{1}{2}} P G^{-\frac{1}{2}}) G^{\frac{1}{2}} \in \mathcal{T}_G.$$

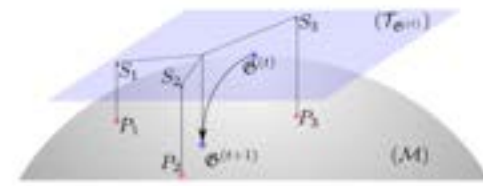
SPDNet architecture



Succession of:

- bilinear layers (BiMap);
- activations (ReEig)

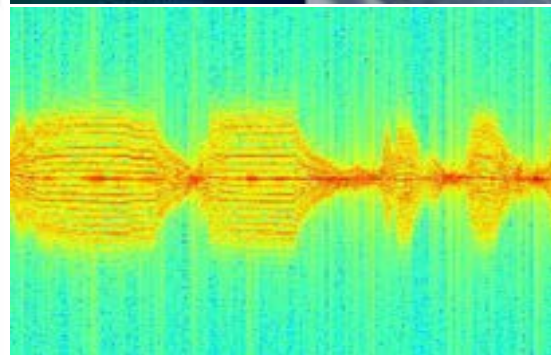
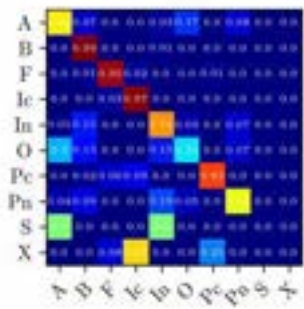
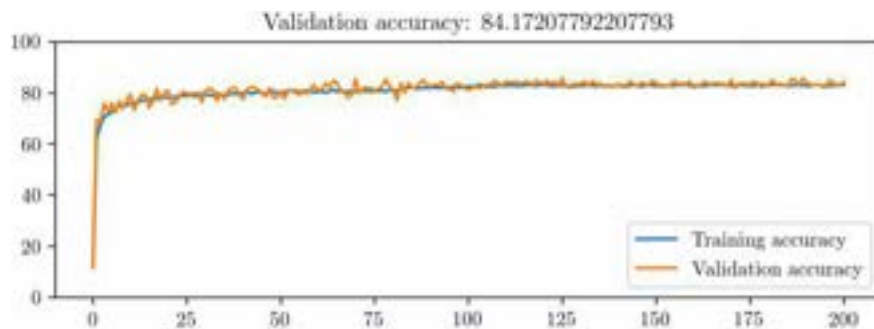
Final manifold: transformed to Euclidean space (LogEig).



Drone Recognition by Radar Micro-Doppler Signature: Results

Validation on NATO Database

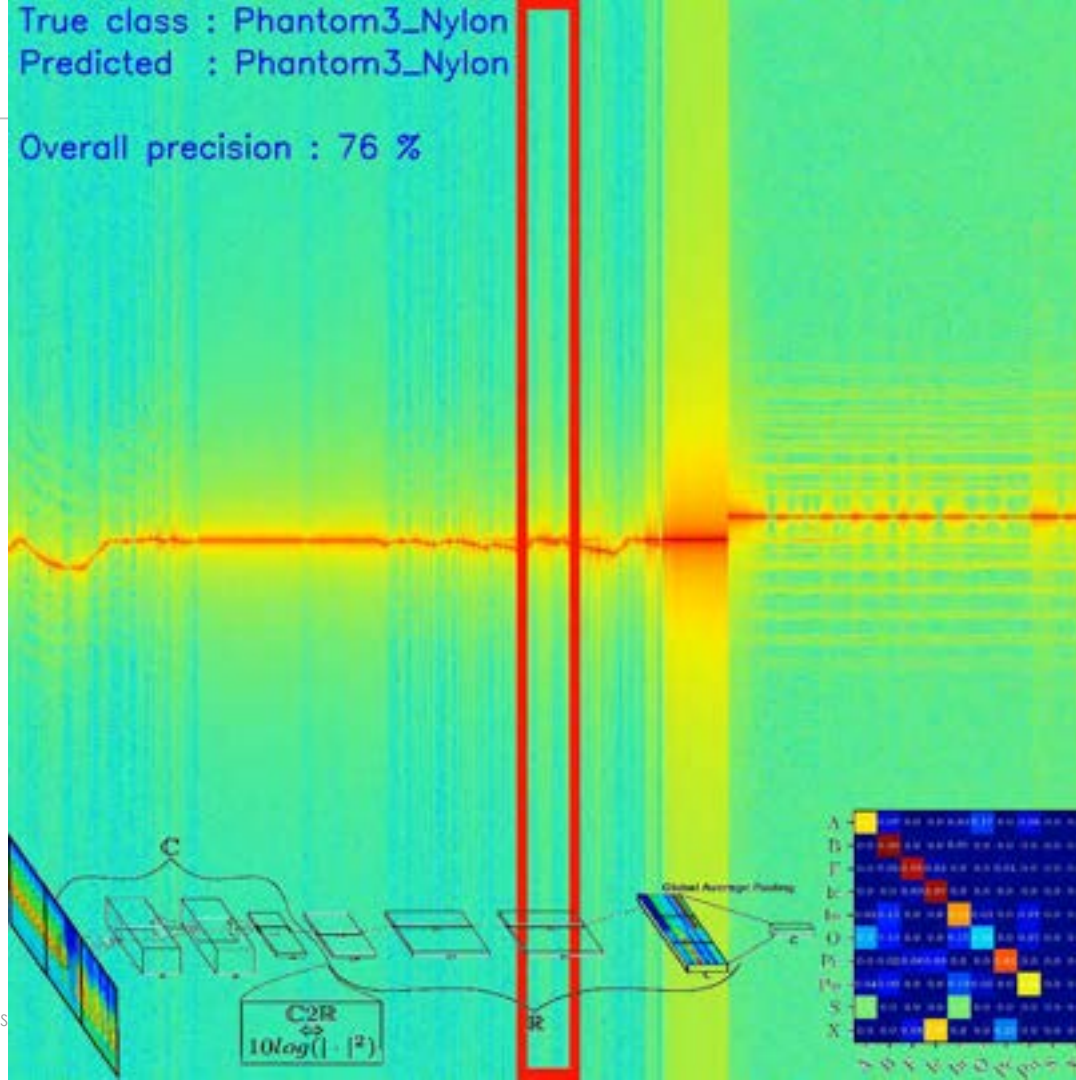
- 10 classes (7 drones and birds)



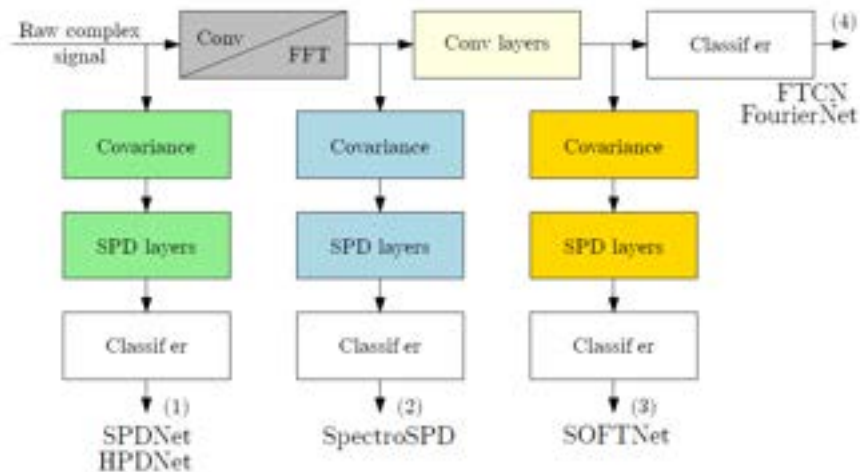
Drone ATDR

True class : Phantom3_Nylon
Predicted : Phantom3_Nylon

Overall precision : 76 %



Performances



Comparing SPD architectures

Performance of SPDNet, DAMNet and SPDNetBN on the NATO dataset

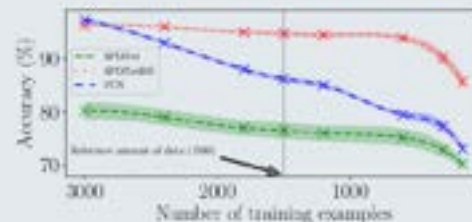
| Model | SPDNet | DAMNet | SPDNetBN |
|-----------------|------------------|------------------|------------------------------------|
| Normalization | - | BarNorm | BatchNorm |
| Accuracy | 72.6% \pm 0.61 | 79.9% \pm 1.19 | 80.3% \pm 0.55 |
| Acc. (10% data) | 69.1% \pm 0.97 | 73.8% \pm 0.25 | 70.2% \pm 1.74 |
| | | | 82.3% \pm 0.80 |
| | | | 77.7% \pm 0.95 |

→ DAMNet and SPDNetBN largely outperform SPDNet.

SOFTNet vs individual models

Performance comparison of first- and second-order models on radar data.

| Train size | 100% | 20% | 5% |
|------------|-----------------|-----------------|-----------------|
| SPDNet | 92.6 \pm 0.54 | 91.5 \pm 0.74 | 88.4 \pm 3.06 |
| HPDNet | 94.4 \pm 0.76 | 91.8 \pm 0.80 | 87.1 \pm 1.11 |
| FTCN | 98.9 \pm 0.44 | 93.4 \pm 1.21 | 84.3 \pm 2.51 |
| FourierNet | 99.4 \pm 0.17 | 96.2 \pm 1.12 | 87.4 \pm 1.94 |
| SpectroSPD | 95.1 \pm 0.49 | 91.9 \pm 0.82 | 84.6 \pm 3.49 |
| SOFTNet | 99.5 \pm 0.16 | 97.2 \pm 0.90 | 93.9 \pm 0.74 |



Performance of models in function of the amount of synthetic radar data: SPD-based models experience a smoother degradation than the deep FTCN.

Fully CNN, SPDNet/HPDNet Architecture and Complex Processing

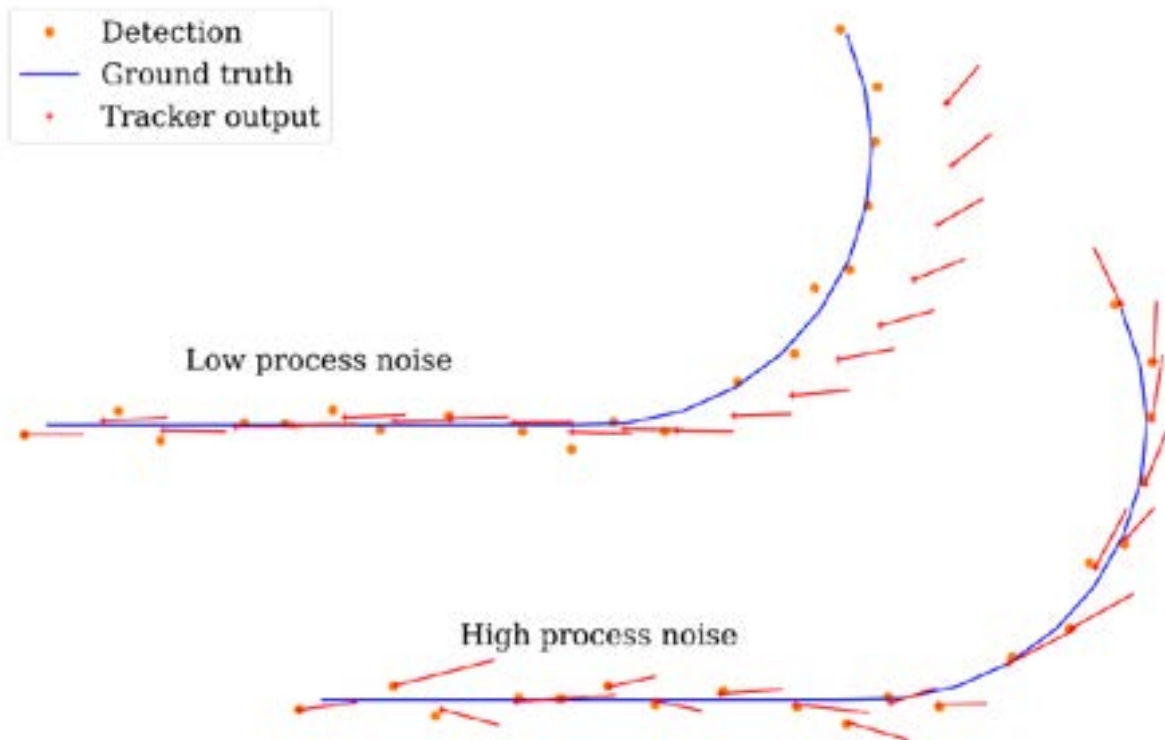
Key takeaways

- Sculpting deep architectures to the data can yield effective, relatively lightweight models.
- SPD-based methods are very lightweight yet not necessarily competitive.
- SPD-based methods seem very robust to lack of data, due to the information-geometric prior.
- Normalization schemes seem to systematically improve SPD networks.
- Combining deep nets with SPD nets allows to build a highly performing and robust learning model, harnessing most input representations of the data.

Advanced Sensor Processing based on Koszul Information Geometry

- Geometric Matrix CFAR/STAP for (very) slow targets detection in clutter
- Complex-Valued CNN & Covariance-Matrix-Valued HPDNet for Micro-Doppler ATDR
- Tracker parameters tuning by Deep Learning for tracking hyper-maneuvering targets
- Multi-Agent Reinforcement Learning for Sensor Resources Management for tracking hyper-maneuvering targets
- Multi-Sensors Collaborative Tracking by Distributed Auctions for tracking in saturating scenario (swarm, fleet of targets, ...)

State of the art: Castella regulation process



An Adaptive Two-Dimensional Kalman Tracking Filter

F.R. Castella
IEEE Transactions on Aerospace
and Electronic Systems, 1980

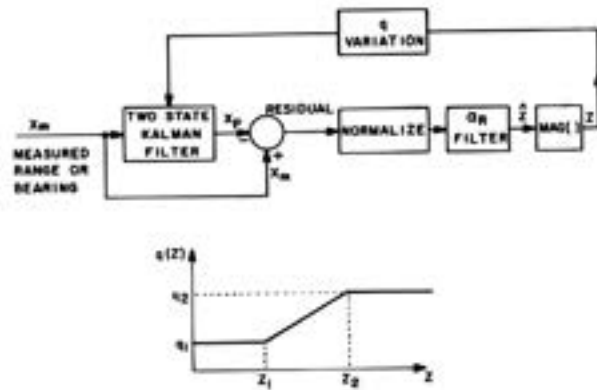
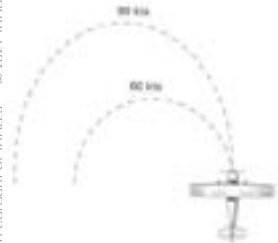
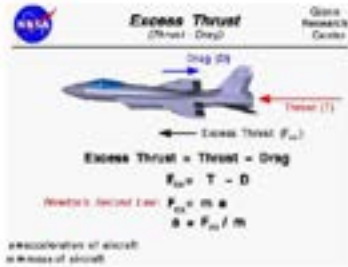


Fig. 1. Tracking for single coordinate.

Flight Mechanics Simulation



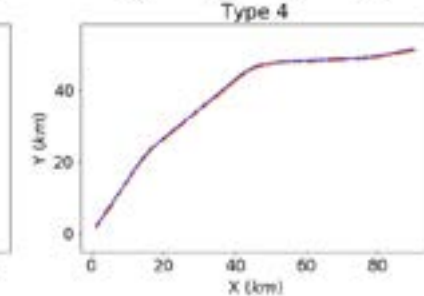
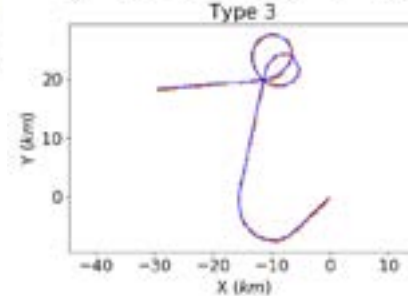
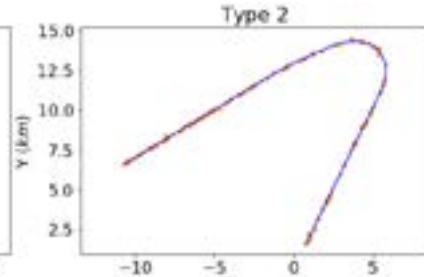
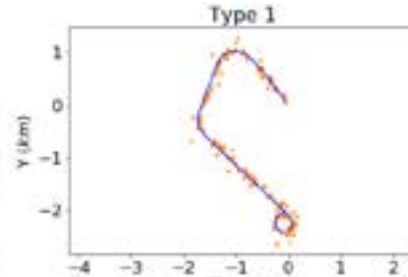
Turn rate



Acceleration



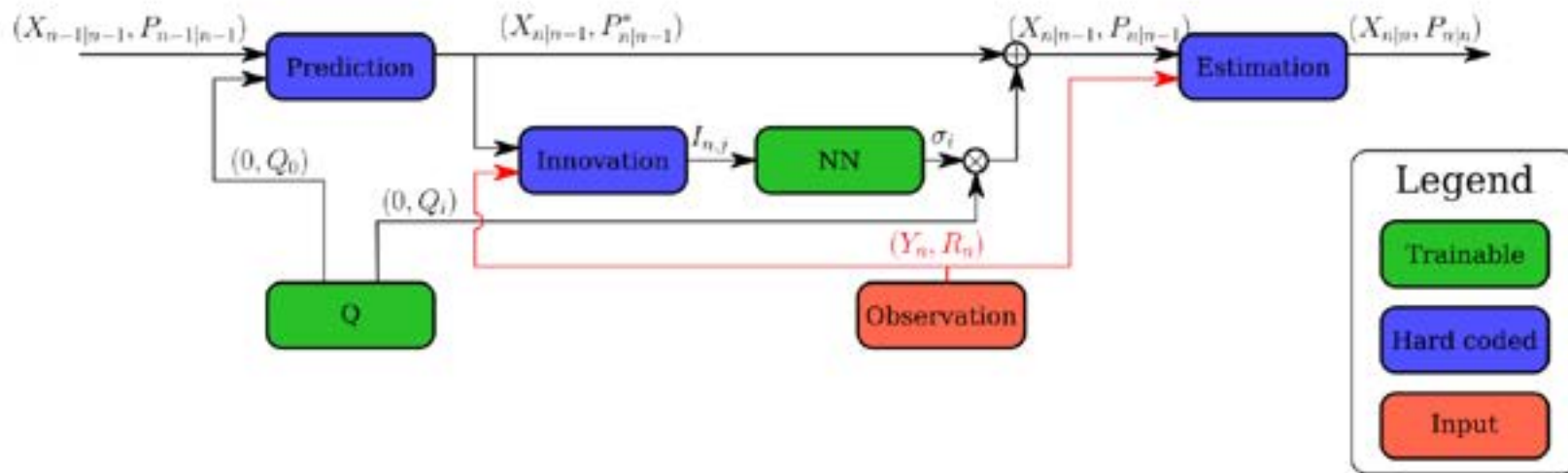
Vertical speed



NNAKF: A Neural Network Adapted Kalman Filter for Target Tracking

S. Jouaber, S. Bonnabel, S. Velasco-Forero, M. Pilté

IEEE ICASSP 2021



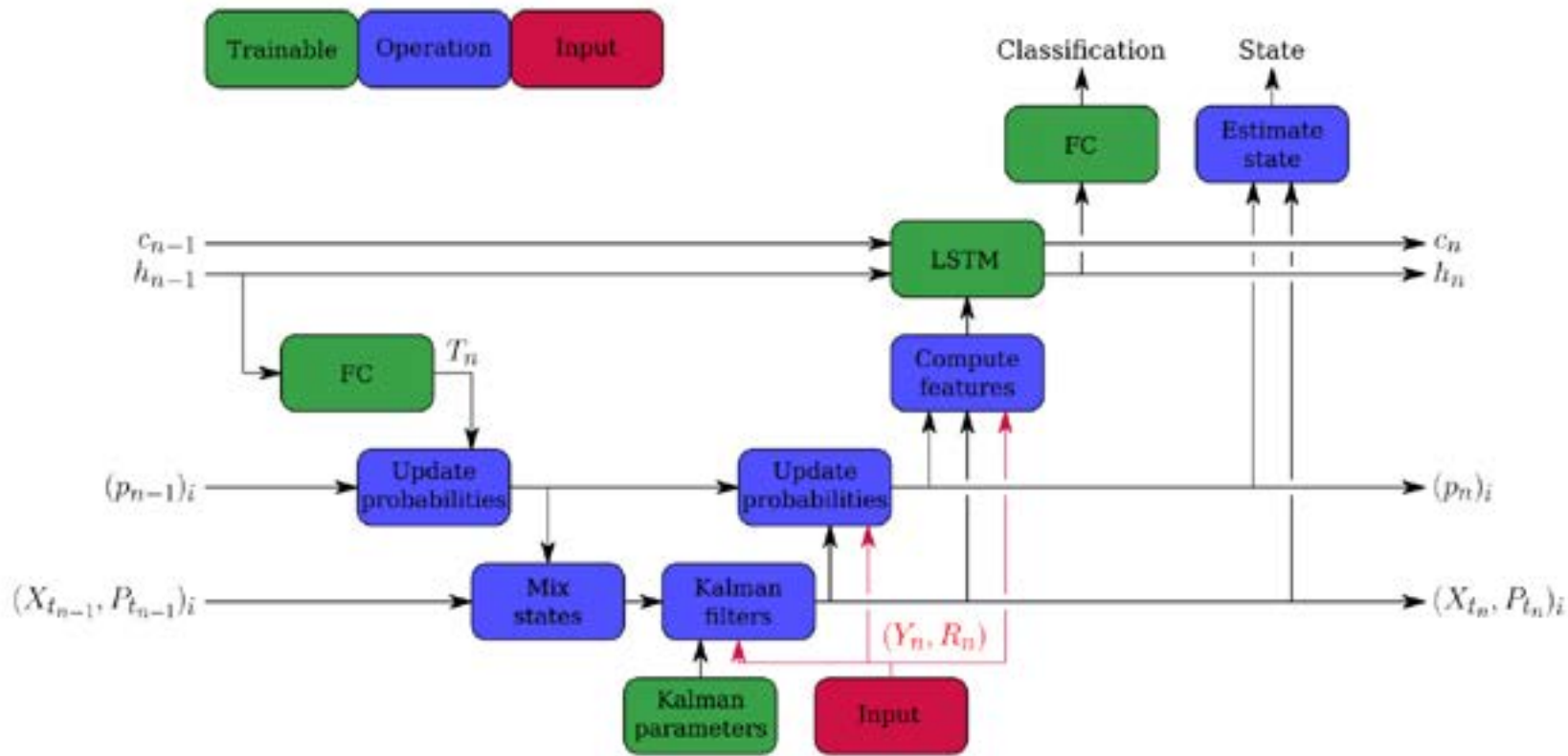
NNAKF versus Castella Performances

| Filter | Position (m) | Altitude (m) | Speed ($m.s^{-1}$) | Heading ($^{\circ}$) |
|----------|--------------|--------------|----------------------|------------------------|
| KF | 35, 91 | 16, 47 | 2, 86 | 5, 09 |
| Castella | 21, 05 | 16, 58 | 1, 86 | 3, 48 |
| NNAKF | 18, 61 | 12, 98 | 1, 48 | 2, 69 |

| Filter | Position | Altitude | Speed | Heading | Average |
|----------|----------|----------|----------|----------|----------|
| Castella | -18, 76% | + 0, 67% | -34, 97% | -31, 63% | -21, 17% |
| NNAKF | -28, 17% | -21, 71% | -48, 25% | -47, 15% | -36, 19% |

This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of THALES - © 2021 THALES. All rights reserved.

Learning Extension for IMM Tracker



Performances Comparison

| Filter | Position (m) | Altitude (m) | Speed ($m.s^{-1}$) | Heading ($^{\circ}$) |
|--------------|------------------|------------------|----------------------|------------------------|
| Oracle | 15, 50 | 8, 73 | 0, 44 | 0, 52 |
| KF | 49, 19 | 35, 49 | 3, 75 | 7, 71 |
| Castella | 44, 33 | 35, 60 | 2, 75 | 6, 10 |
| NNAKF | 41, 89 | 32, 00 | 2, 37 | 5, 31 |
| IMM 20 CT | 37, 25 | 28, 61 | 1, 75 | 4, 21 |
| IMM+NN 20 CT | 35, 58 | 27, 20 | 1, 63 | 3, 56 |

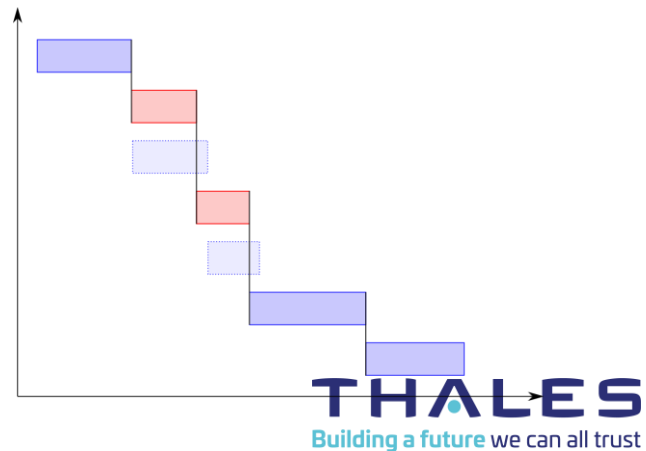
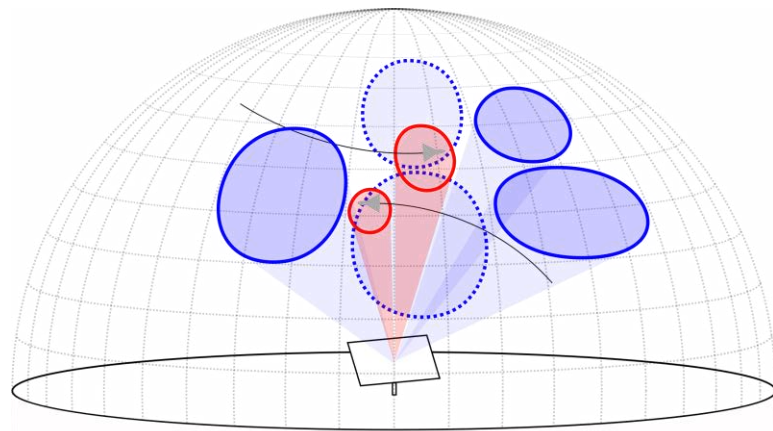
Advanced Sensor Processing based on Koszul Information Geometry

- Geometric Matrix CFAR/STAP for (very) slow targets detection in clutter
- Complex-Valued CNN & Covariance-Matrix-Valued HPDNet for Micro-Doppler ATDR
- Tracker parameters tuning by Deep Learning for tracking hyper-maneuvering targets
- Multi-Agent Reinforcement Learning for Sensor Resources Management for tracking hyper-maneuvering targets
- Multi-Sensors Collaborative Tracking by Distributed Auctions for tracking in saturating scenario (swarm, fleet of targets, ...)

Context and objectives

Multi-function radars (MFR)

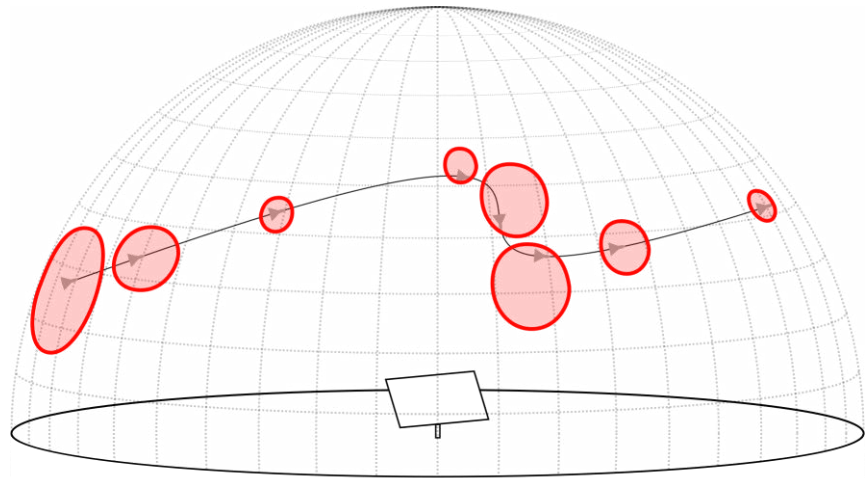
- Must optimize the performance of many tasks in parallel and in real time
 - **Search task**: maintain coverage of surveillance space to maximize detection probability
 - **Tracking tasks**: estimate and predict each target's state with a given level of precision
 - Tasks issued by a command & control center: missile uplink, kill assessment, etc.
- Tasks have a common limiting resource: antenna time budget
 - Dwells must be scheduled on a common timeline
 - In overload situations, the lower-priority dwells must be delayed/abandoned due to time constraints



Context and objectives

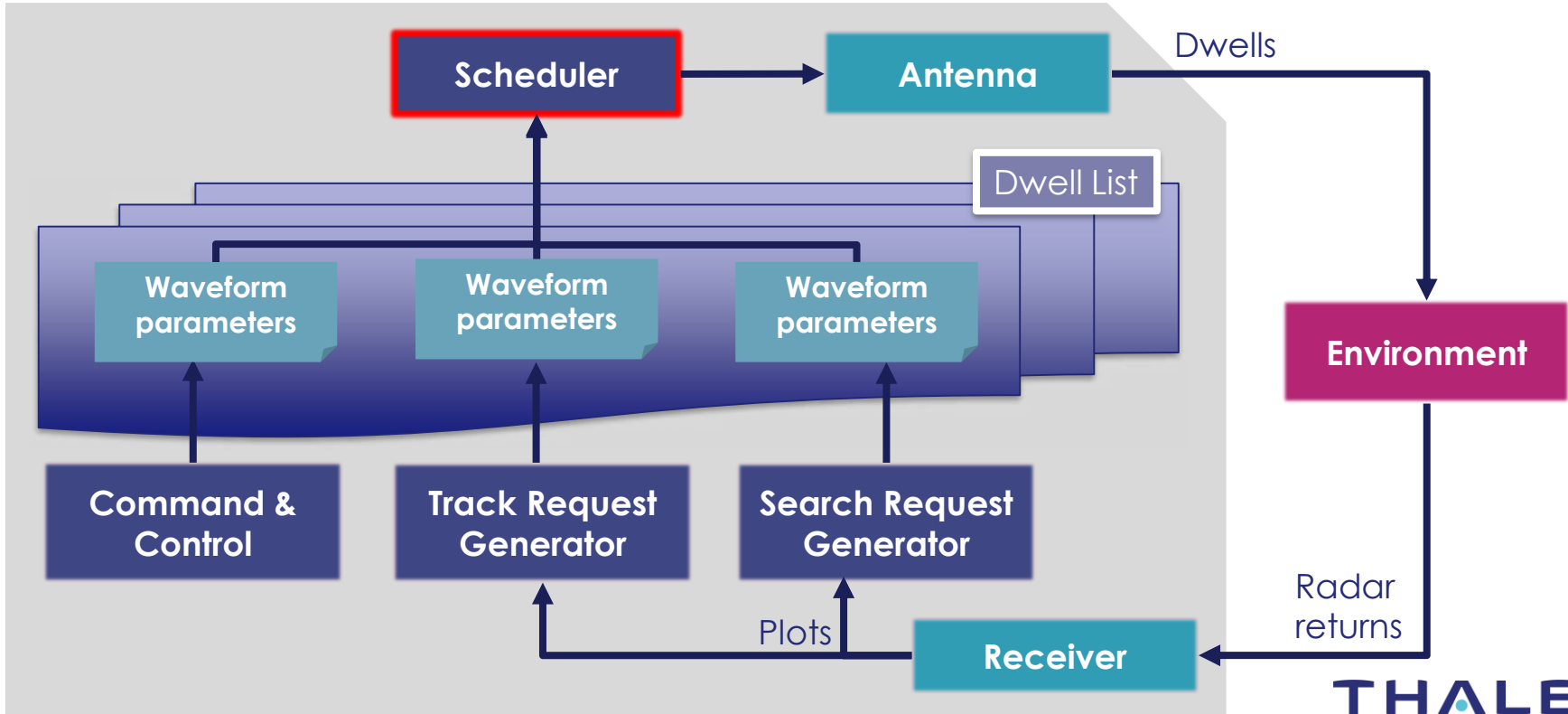
Objectives

- Enhance tracking with adaptive revisit interval (RI) and adaptive beamwidth (BW)
 - Made possible in recent years by digital beamforming in phased-array antennas
 - Could help tracking difficult targets (hypermaneuvering, hypersonic, reactive)
 - Currently one BW for all dwells, one fixed RI for each track
- Handle overload more efficiently
 - More likely to happen due to swarm attacks
 - Current approaches using rule-based heuristics can be hard to adapt
- Coordinate search and tracking
 - Could use search for reacquisition of lost targets when possible
 - Instead of using time-consuming reacquisition patterns
- See how reinforcement learning (RL) can help with all of this



System architecture

Multi-function radar

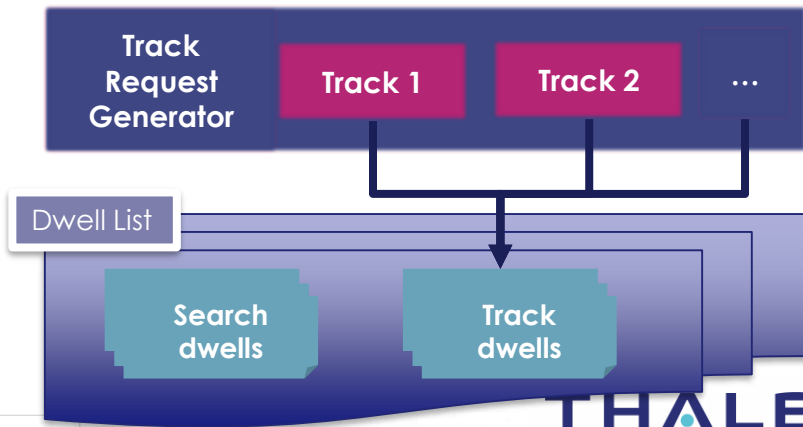


MFR as a multi-agent system

- Multiple conflicting tasks within the radar point to a multi-agent structure
 - One agent per tracked target
 - Each track must decide which dwells are adequate for its target's position/velocity estimation
 - In this sense, each track constitutes an independent subproblem
 - Search is considered an artefact (i.e. not an agent, but a resource)
 - No autonomy: search dwell maps are precomputed and only represent a fixed list

➤ Advantages:

- Explainability
- Extensibility (to radar networks, etc.)
- Simpler to integrate domain constraints
- Simpler optimization
 - Tracks = instances of same problem

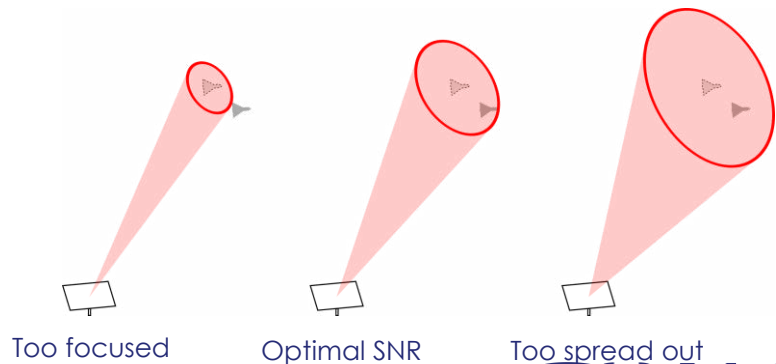
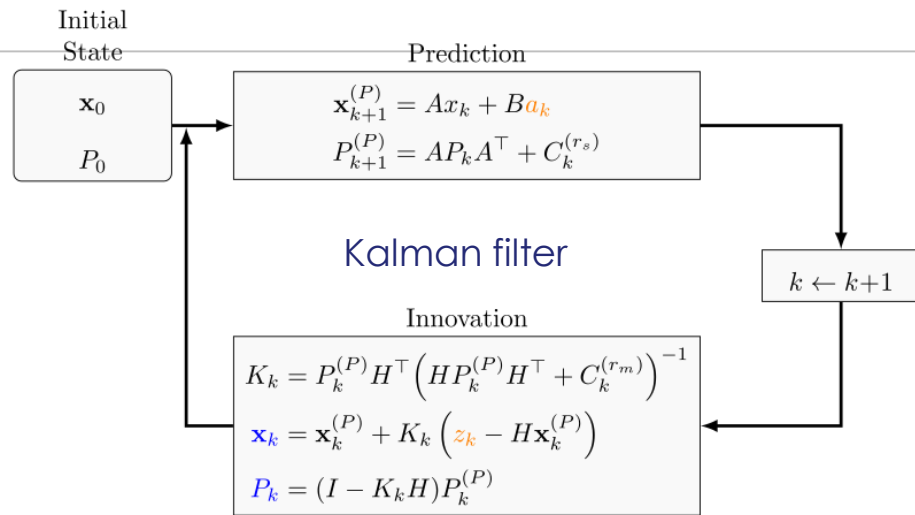


OPEN

System architecture

Objective of a tracking agent

- Parameters to optimize: revisit interval (RI) and beamwidth (BW)
- Objective: minimize the target's state estimate error covariance (EEC) P_k , i.e. the uncertainty over the target's true state x_k^* , as evaluated by the Kalman filter
- EEC increases predictably as time since last revisit increases
- The optimal BW (in terms of SNR) should depend on the prediction error $x_k^* - x_k^{(P)}$
- No model to predict the effect of a dwell on EEC, must be learned



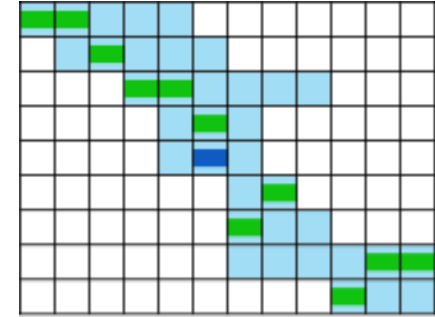
Radar task scheduling

Formalization

➤ For each dwell $i \in I$:

- We know the earliest and latest possible execution start times T_s^i and T_{dr}^i , the ideal execution time T_{du}^i , and the shortest and longest possible durations L_{min}^i and L_{max}^i
- As a measure of priority, via threat model – EEC aggregation, we can define :
 - A drop cost C_{dr}^i , based on predicted EEC increase if this dwell is not executed
 - A duration cost C_{le}^i , based on how much EEC would degrade if dwell was shortened (narrower beamwidth)
 - A delay cost C_{de}^i , to penalize divergence from T_{du}^i
- We want to know whether to schedule the dwell ($s^i = 1$), and if so at what execution time t^i , or to drop it ($d^i = 1$)

$$\min C_P(s) = \sum_{i \in I} s^i |t^i - T_{du}^i| C_{de}^i + s^i (L_{max}^i - l^i) C_{le}^i + d^i C_{dr}^i$$
$$\text{s.t.} \begin{cases} T_s^i \leq t^i \leq T_{dr}^i \quad \forall i \in I \\ t^i + l^i \leq T_{max} \quad \forall i \in I \\ t^i + l^i \leq t^j \oplus t^j + L^j \leq t^i \text{ if } s^i = s^j = 1, \forall (i < j) \in I^2 \\ L_{min}^i \leq l^i \leq L_{max}^i \quad \forall i \in I \\ t^i, l^i \in \mathbb{R}^+ \quad \forall i \in I \\ s^i, d^i \in \{0, 1\} \text{ with } s^i = 1 - d^i, \forall i \in I \end{cases}$$



Resolution methods

- Heuristics (EST, EDF): fast, but do not take priorities into account
- Mixed integer programming: returns the optimal solution – given enough time...
 - Complete formulation:

$$\begin{aligned} \min \quad & \sum_{i \in I} l_{dc}^i C_{dc}^i + (1 - s^i) C_{dr}^i \\ \text{s.t.} \quad & \left\{ \begin{array}{l} t^i \geq T_s^i \\ t^i \leq T_{dr}^i \\ t^i + L^i \leq T_{max} \\ l_{dc}^i \geq t^i - T_{du}^i \\ l_{dc}^i \geq T_{du}^i - t^i \end{array} \right\} \forall i \in I \\ & \left\{ \begin{array}{l} n^{ij} \leq 2 - s^i - s^j \\ t^i + L^i \leq t^j + M(n^{ij} + x^{ij}) \\ t^j + L^j \leq t^i + M(n^{ij} + 1 - x^{ij}) \\ t^i, l_{dc}^i \in \mathbb{R}^+ \forall i \in I \\ \boxed{s^i \in \{0, 1\} \forall i \in I} \\ \boxed{n^{ij}, x^{ij} \in \{0, 1\} \forall (i < j) \in I} \end{array} \right\} \forall (i < j) \in I \end{aligned}$$

This problem is NP-hard!

Monte Carlo Tree Search

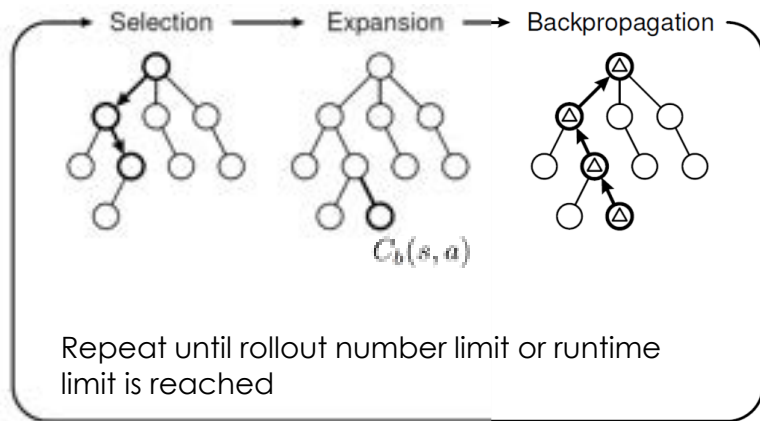
➤ Model-based method that explores a search tree

- Root = empty schedule
- Leaf = complete schedule with cost C_b
- Backpropagation: $C_b(s, a) = \min_{a'} C_b(s', a')$

- Selection: $a_t = \operatorname{argmax}_a (U(s_t, a))$, $U(s, a) \propto \frac{P(s, a)}{C_b(s, a)^\tau (1 + N(s, a))}$, $a \in \{1, \dots, M\}$

→ **priori** (for example: HCLR heuristic)

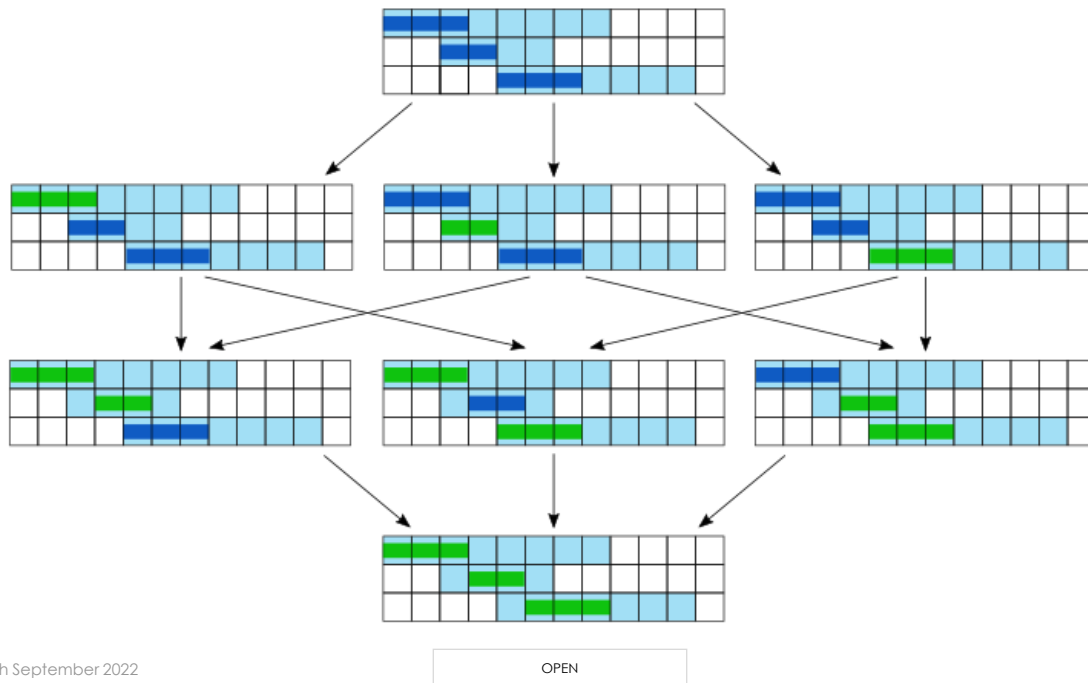
→ number of visits



Resulting policy: $\pi(s) = \operatorname{argmin}_{a'} C_b(s, a')$

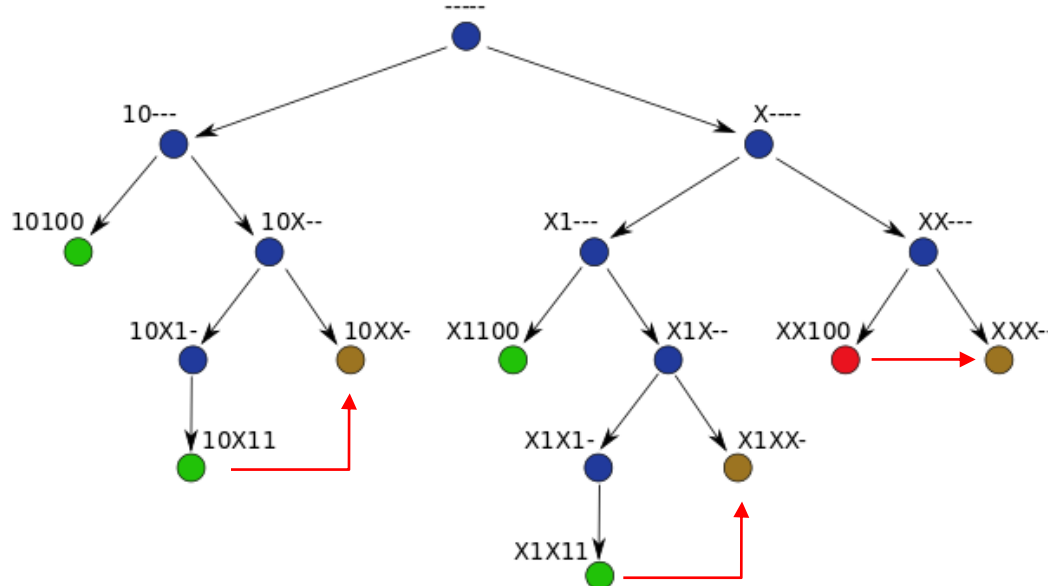
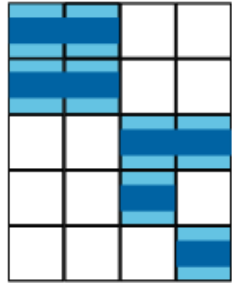
Monte Carlo Tree Search

- Problem : with this transition function, the same complete schedule can be reached through different paths



Monte Carlo Tree Search

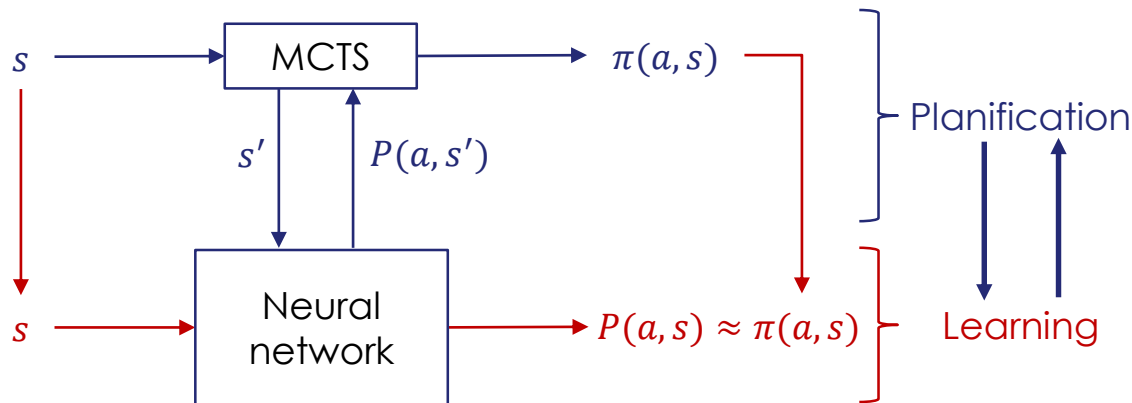
- Solution: structure our search tree as a Branch-and-Bound tree
 - By allowing already visited actions, each leaf can only be reached once
 - Branches can be pruned based on existing solutions
 - If a leaf isn't a complete schedule, its derived branches can be pruned



Reinforcement learning for MCTS

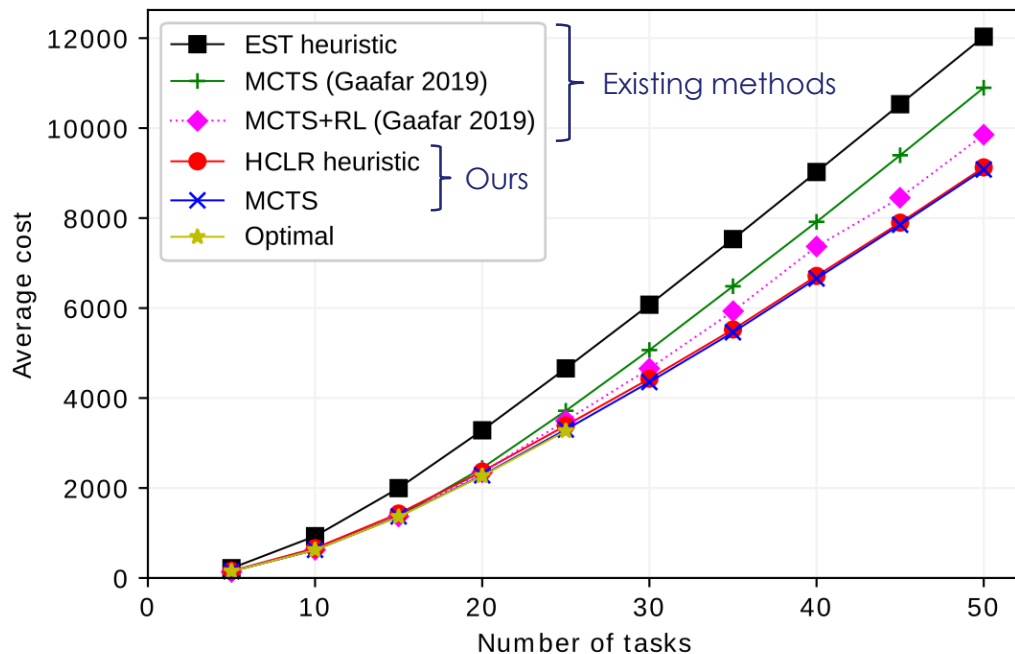
> AlphaZero

- Learn the MCTS prior with a neural network
- Mutual improvement: the neural network guides MCTS, whose results correct the network's predictions



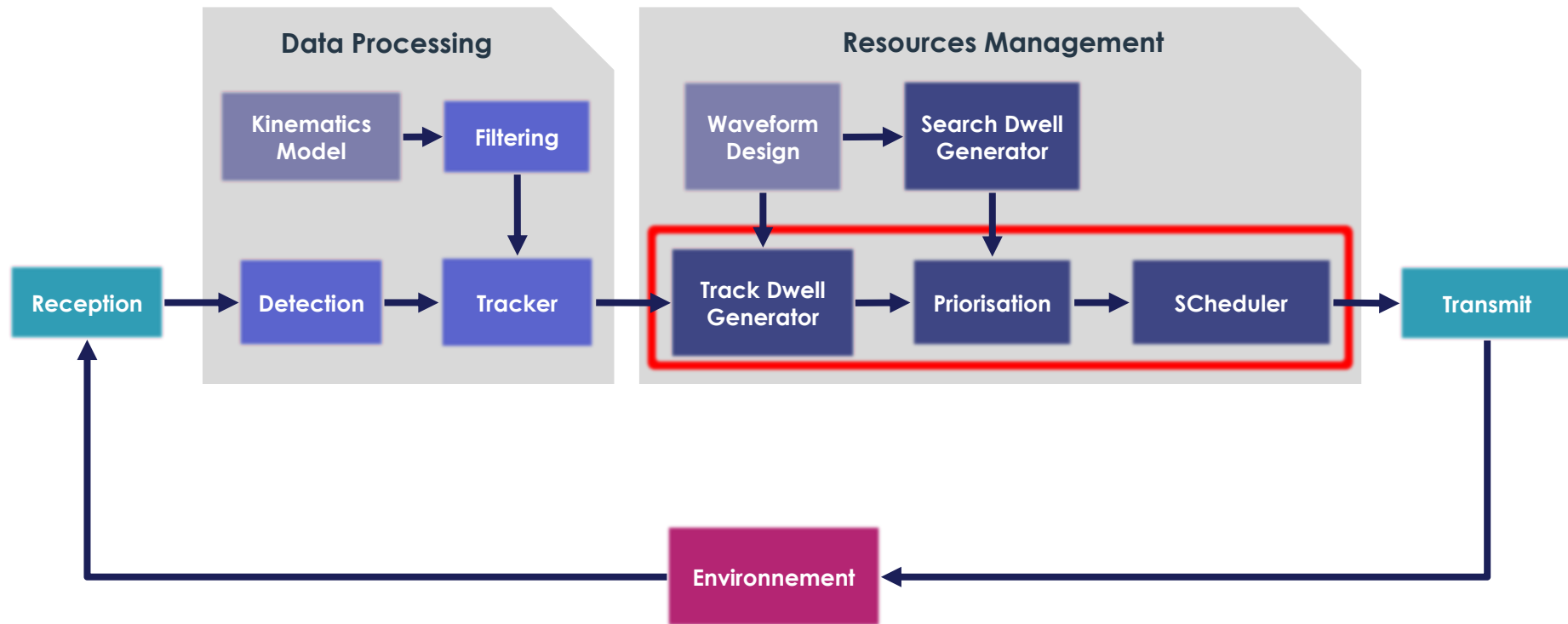
Results

- Experiments on a synthetic benchmark
- Choosing an adequate transition function has more impact on performance than choosing the actual decision algorithm
 - Our heuristic finds near-optimal solutions in a fraction of the time required for tree-search methods



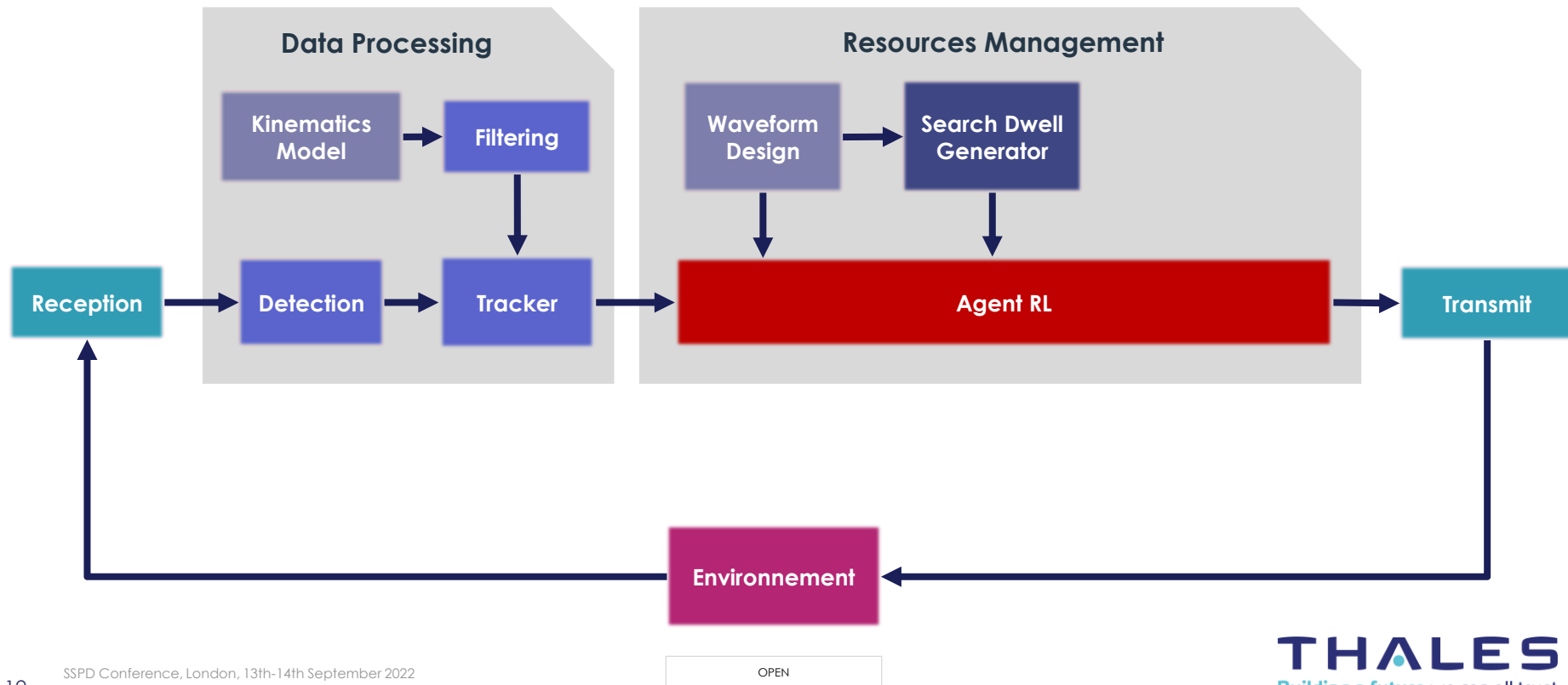
MFR Radar Resources Management

Functional Architecture



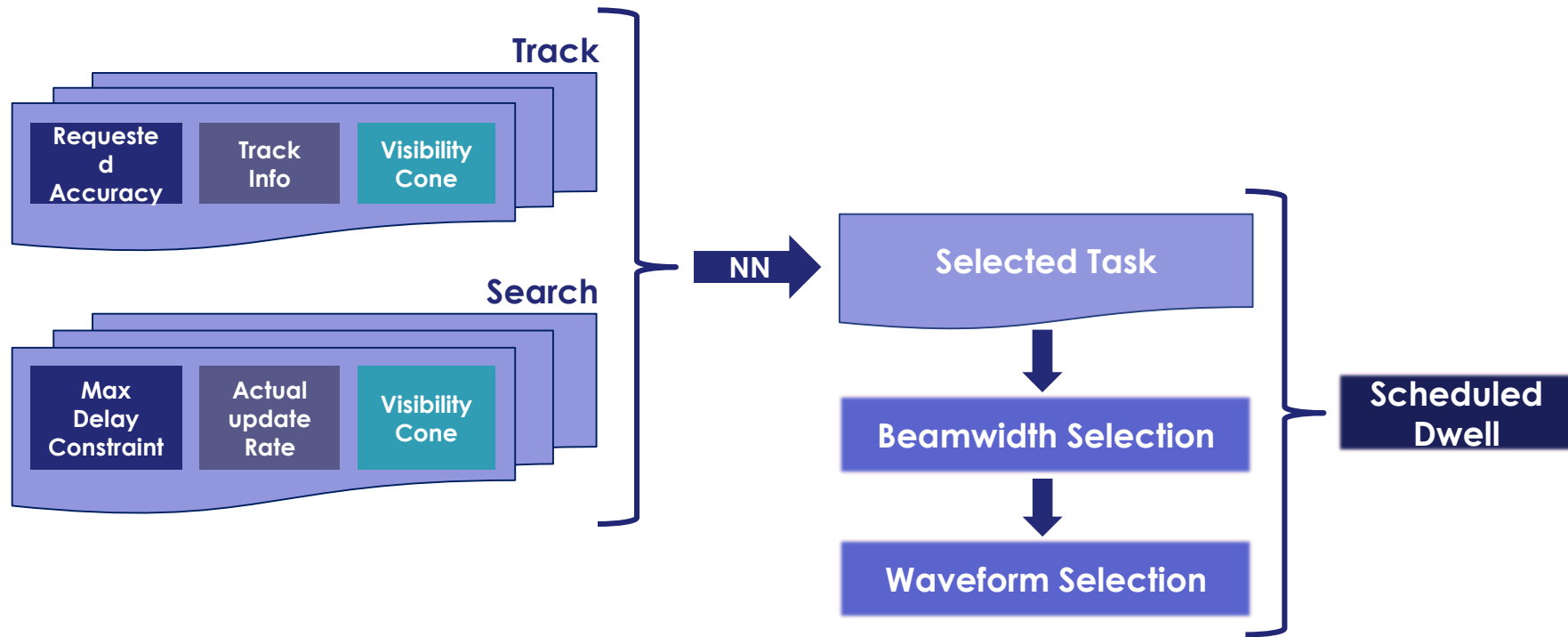
Multi-Agent RL for MFR Radar Management

Functional Architecture



RL Actions selection for Beam update rate and Beamwidth

RL Agent Structure



Advanced Sensor Processing based on Koszul Information Geometry

- Geometric Matrix CFAR/STAP for (very) slow targets detection in clutter
- Complex-Valued CNN & Covariance-Matrix-Valued HPDNet for Micro-Doppler ATDR
- Tracker parameters tuning by Deep Learning for tracking hyper-maneuvering targets
- Multi-Agent Reinforcement Learning for Sensor Resources Management for tracking hyper-maneuvering targets
- Multi-Sensors Collaborative Tracking by Distributed Auctions for tracking in saturating scenario (swarm, fleet of targets, ...)

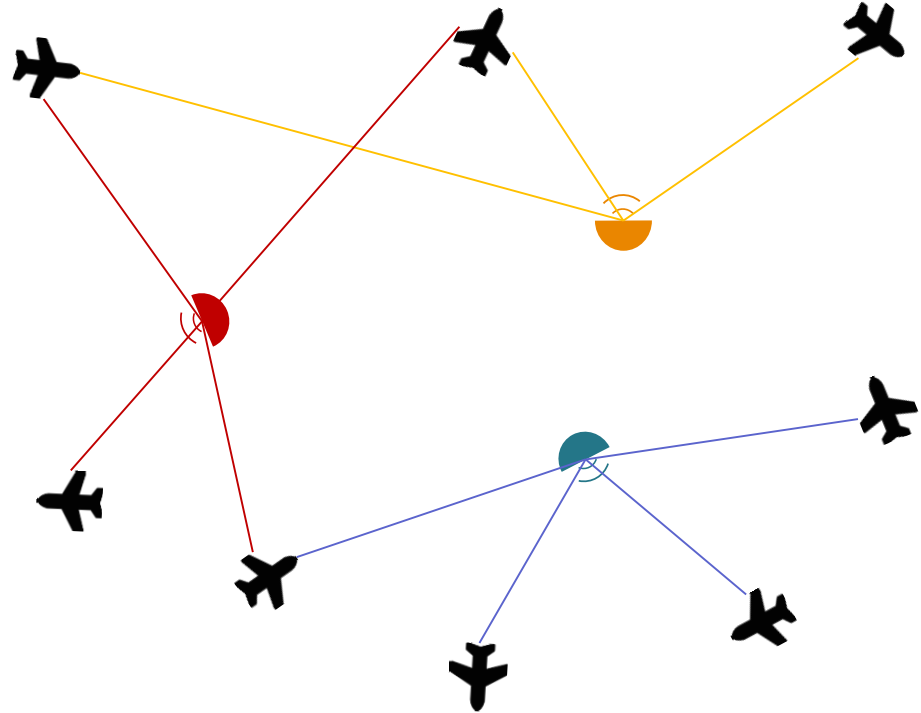
Context & Problem statement

Context

- Active-tracking with cooperation of up to 2 radars
- Load Balancing
- Resilience

Problem statement

- Given a radar configuration, allocate at best resources
- When coupling radars, ensure to provide additional information if possible
- Centralized vs Decentralized coordination
- Standby phase ignored



Radar & Target Models

Directed radar

Information:

$$\text{Radar load} = \frac{D_i}{T_i}$$

$$T_i = cst$$

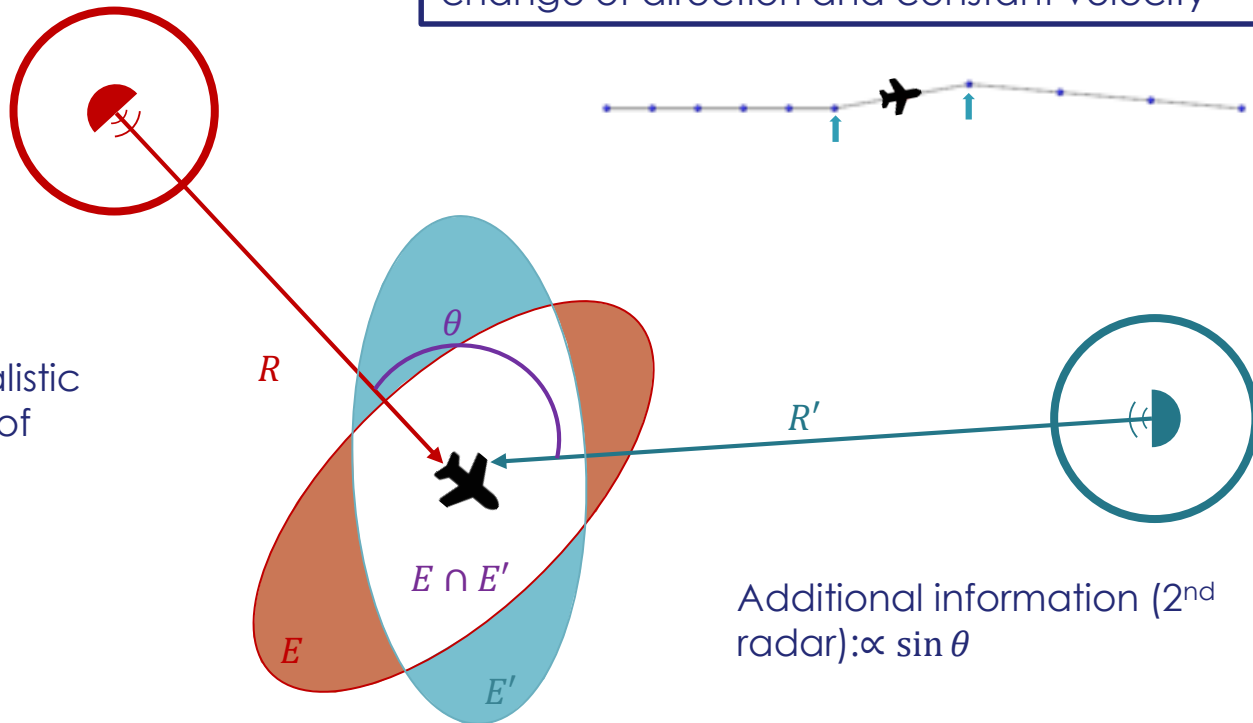
$$D_i \propto R$$

$$SNR = cst$$

Overload limit unrealistic
(goal: explore limits of
the radars)

Tracking uncertainty
determined by a
Kalman Filter

Targets have a straight-path model with occasional
change of direction and constant velocity



Additional information (2nd
radar): $\propto \sin \theta$

OPEN

Centralized Optimization Problem Formulation

Mono-Sensor Allocation : $|\mathcal{I}| + |\mathcal{J}|$ constraints.

$$(P1) : \begin{cases} \max \sum_{i,j} c_{ij} \cdot x_{ij} \\ \text{s.t.:} \\ \sum_i x_{ij} \leq 1, \forall j \in \mathcal{J} & (C1) \\ \sum_j \gamma_{ij} \cdot x_{ij} \leq L_{ti}, \forall i \in \mathcal{I} & (L) \\ x_{ij} \in \{0, 1\}, \forall (i, j) \in \mathcal{I} \times \mathcal{J} \end{cases}$$

Sets :

- \mathcal{I} : Set of cardinality radars $|\mathcal{I}| = N_u$
- \mathcal{J} : Set of cardinality tasks $|\mathcal{J}| = N_t$

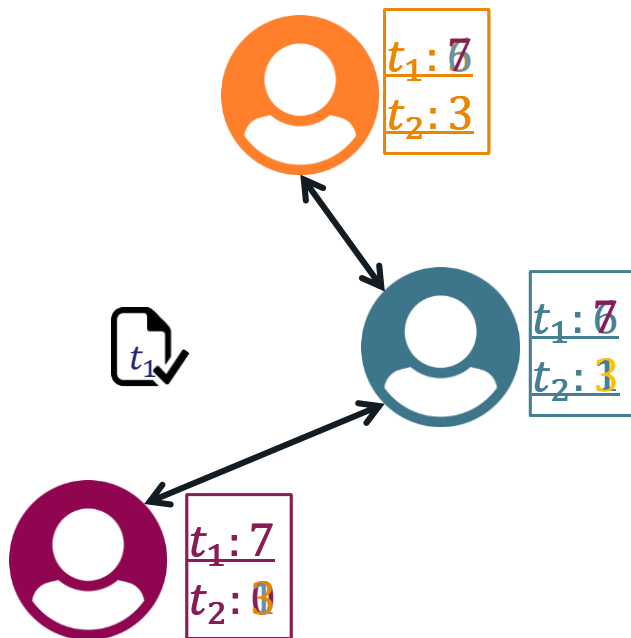
2-Sensors Allocation: $|\mathcal{I}|^2 \cdot |\mathcal{J}| + |\mathcal{J}| + |\mathcal{I}|$ constraints

$$(P2) : \begin{cases} \max \sum_{i,j,k} c_{ikj} \cdot w_{ikj} \\ \text{s.t.:} \\ w_{ikj} = x_{Mij} \wedge x_{Okj}, \forall (i, k) \in \mathcal{I}^2, \\ \forall j \in \mathcal{J} (A_{ikj}) \\ \sum_i w_{ij} \leq 1, \forall j \in \mathcal{J} & (C2) \\ \sum_j \gamma_{ij} \cdot (x_{Mij} + x_{Okj} - w_{iij}) \leq L_{ti}, \\ \forall i \in \mathcal{I} & (L) \\ (x_{Mij}, x_{Okj}) \in \{0, 1\}^2, \forall (i, j) \in \mathcal{I} \times \mathcal{J} \\ w_{ikj} \in \{0, 1\}, \forall (i, k) \in \mathcal{I}^2, \forall j \in \mathcal{J} \end{cases}$$

CBBA: A Decentralized Sealed Auction Algorithm



Consensus reached !



Consensus Based Auction Algorithm, a 2 phase algorithm:

- Bidding phase
- Consensus phase

Bidding phase:

- Each Agent evaluates its bid for given tasks
- Bid based on utility

Consensus phase:

- Send bid to neighbors
- Update maximal utility from others knowledge
- Send updated information until consensus is reached

Adaptation of CBBA to radars and dynamism

Bids are made for a particular time :

- Data synchronization: attach timestamp to each knowledge
- Data update is performed according to « Update Rules Table »

Diminishing Marginal Gain constraint

- Biased utility to ensure c

$$c_{ij}(b_i) \geq c_{ij}(b_i \oplus b) \text{ (DMG)} \rightarrow c_{ij}^{\text{CBBA}} = \frac{c_{ij}}{|b_i|}$$

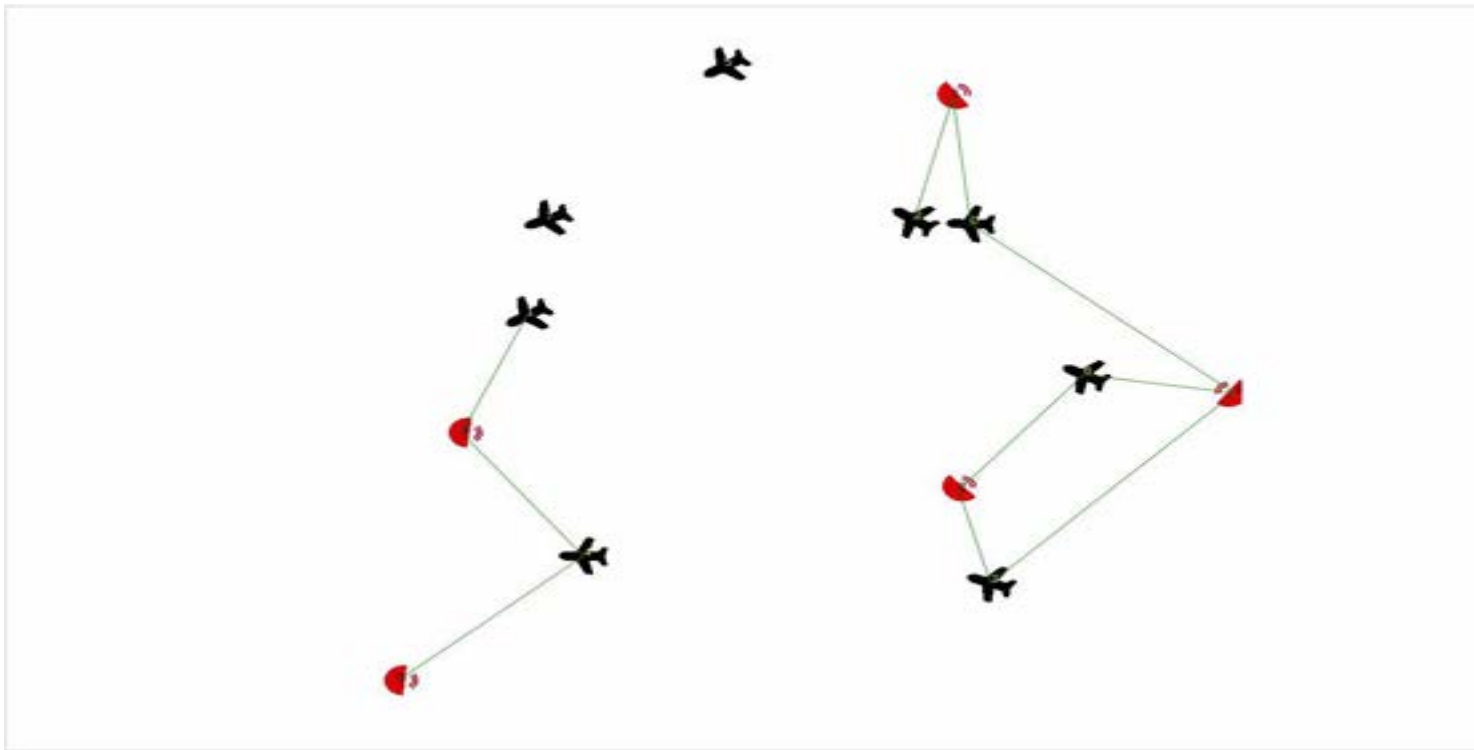
- Adding a task is always less profitable than to be dedicated to a given set of task
- Prevent a radar to track all targets if it is possible → Load Balancing

| The transmitter (radar k) thinks the winner is: | The recipient (radar i) thinks the winner is: | Action taken by i |
|--|--|--|
| k | i | If $y_{ij} > y_j \rightarrow \text{update}$ |
| | k | update |
| | $m \notin \{i, k\}$ | If $y_{ij} > y_j$ and $s_{km} > s_{im} \rightarrow \text{update}$ |
| | none | update |
| i | i | leave unchanged |
| | k | Reset |
| | $m \notin \{i, k\}$ | if $s_{km} > s_{im}$ reset |
| | none | leave unchanged |
| $m \notin \{i, k\}$ | i | If $y_{ij} > y_j$ and $s_{km} > s_{im} \rightarrow \text{update}$ |
| | k | if $s_{km} > s_{im} \rightarrow \text{update}$ |
| | | otherwise $\rightarrow \text{reset}$ |
| | m | if $s_{km} > s_{im} \rightarrow \text{update}$ |
| | $n \notin \{i, k, m\}$ | if $s_{kn} > s_{in}$ and $s_{km} > s_{im} \rightarrow \text{update}$ |
| | | if $s_{km} > s_{im}$ and $y_{ij} > y_j \rightarrow \text{update}$ |
| | | if $s_{kn} > s_{in}$ and $s_{km} > s_{im} \rightarrow \text{reset}$ |
| none | none | if $s_{km} > s_{im} \rightarrow \text{update}$ |
| | i | leave unchanged |
| | k | update |
| | $m \notin \{i, k\}$ | if $s_{km} > s_{im} \rightarrow \text{update}$ |
| | none | leave unchanged |

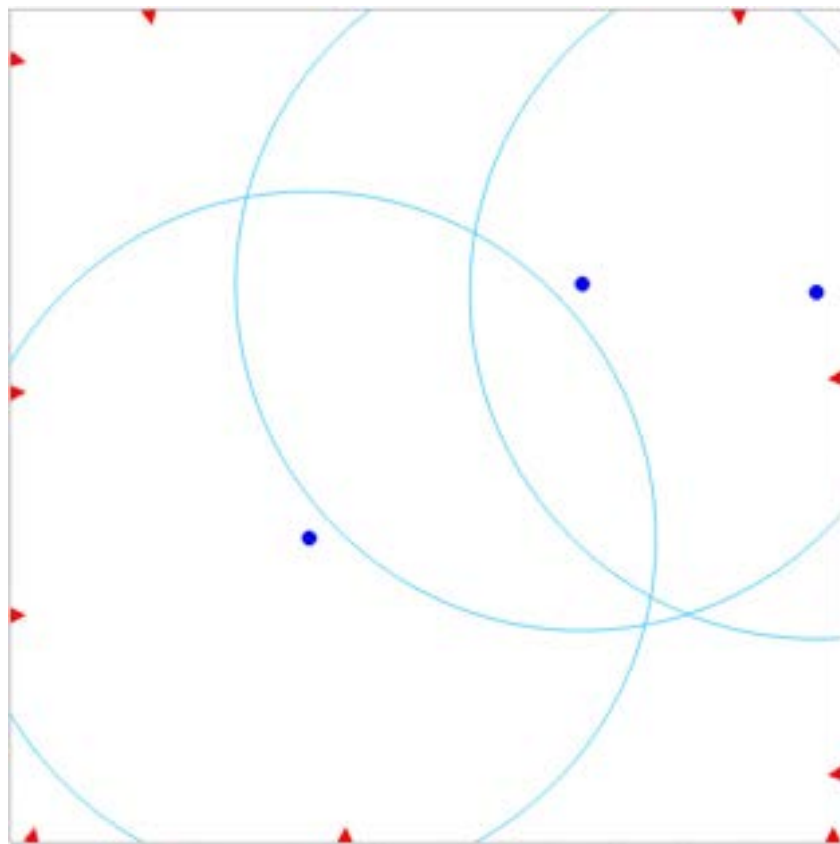
Algorithm operates in closed loop and re-executed periodically :

- 1st phase: CBBA bid step for Main Allocation
 - Utility is proportional the surface of the Kalman uncertainty ellipse
- 2nd phase: CBBA consensus step for Main Allocation
 - Radars update their knowledge for Main Allocation
 - Start/continue tracking on targets with winning bids (as Main)
 - Eventually release optional target tracking
- 3rd phase: CBBA bid step for Optional Allocation
 - Utility is proportional the surface of the intersection the Kalman uncertainty ellipses
- 4th phase: CBBA consensus step for Optional Allocation
 - Radars update their knowledge for Optional Allocation
 - Start/continue tracking on targets with winning bids (as Optional)

Demo



task allocation for radar network (TRT work with Distributed Auctions)



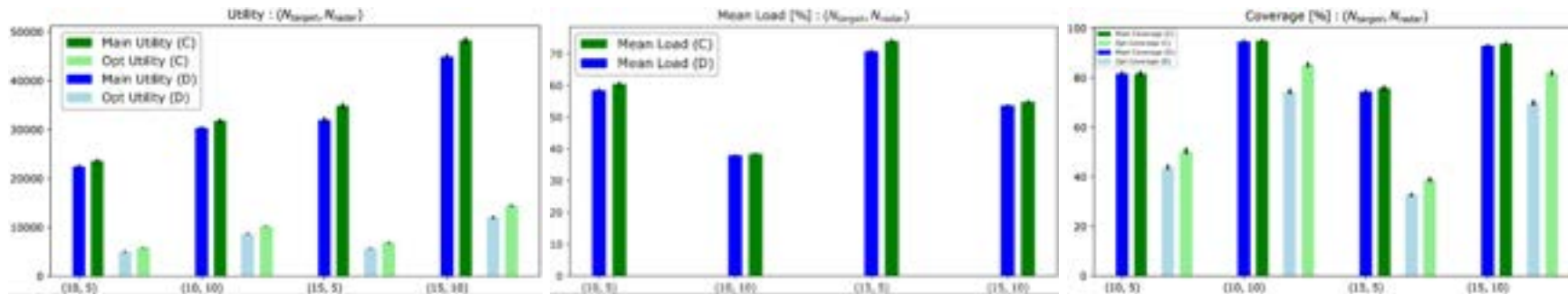
Results : Centralized vs Decentralized approaches

Performance of Decentralized approach higher than theoretical 50% performance guarantee (greedy algorithm)

Load more evenly distributed for Decentralized approach

Resilience aspect

➤ From « star-like » communication network (centralized) to ad-hoc



Modern Sensor Processing based on Symplectic Model of Information

- Lie Group Based Equivariant GCNN for Adaptive Doppler Clutter Map
- Lie Group Based Frenet-Serret IEKF (Invariant Extended Kalman Filter) for tracking hyper-maneuvering targets
- Lie Group Based Target Recognition on Kinematics for Drone/Birds Classification
- Souriau Symplectic Model of Information for Lie Group Statistics and Machine Learning
 - Entropy as Casimir Function in Coadjoint Representation
 - Koszul-Fisher Metric on Lie Group
 - Covariant Maximum Entropy Density (Gauss Density) on Lie Group
 - Lie Groups Machine Learning

Motivations

Conventional Deep Learning weaknesses:

- Lack of **robustness**
- Do not use the **native geometry** of inputs

➔ **Geometric-Informed Algorithms** as remediation

➔ Emerging technologies with already **successful applications** (e.g., AlphaFold@DeepMind)



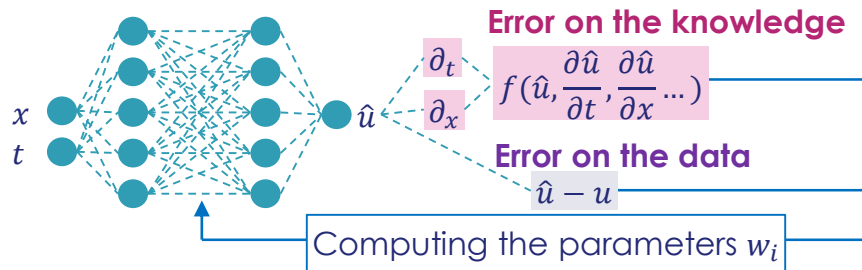
Deep learning requires large amounts of training data and numerical simulation requires a perfect knowledge of the underlying physics:

➔ **Physics-Informed NN** allows for:

- More frugal learning
- Discovering hidden physics
- Reuse the knowledge of similar tasks



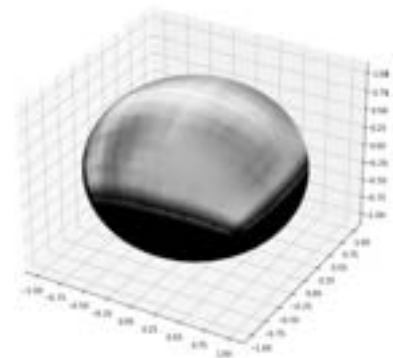
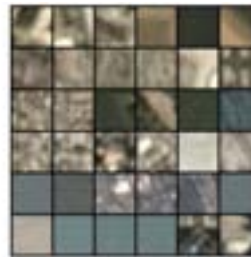
Physics Informed Neural Network



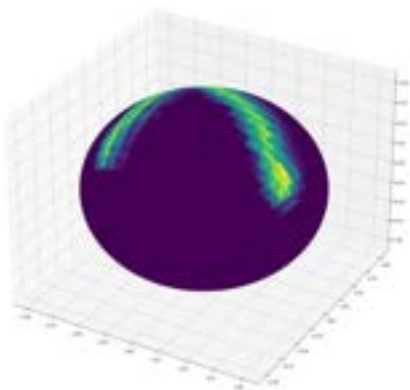
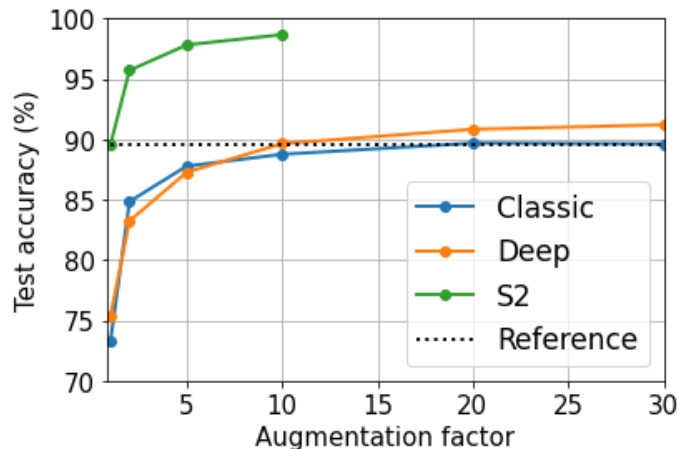
Geometric-Informed NN: Technological progress

Equivariant Neural Networks

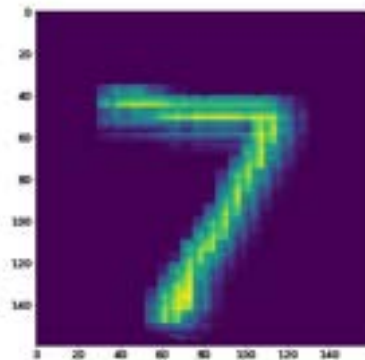
- Create relevant **synthetic datasets**
- **Compare** the use of **equivariance** mechanisms with **data augmentation**
- Design **new convolution operators** on hyperbolic spaces (Poincaré Disk)



on



MNIST image on the sphere



Projected image

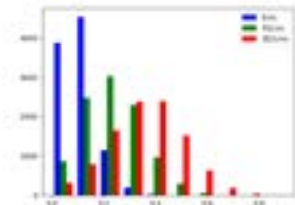
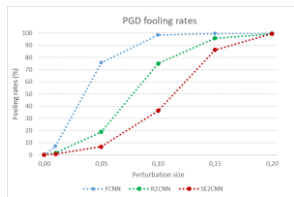
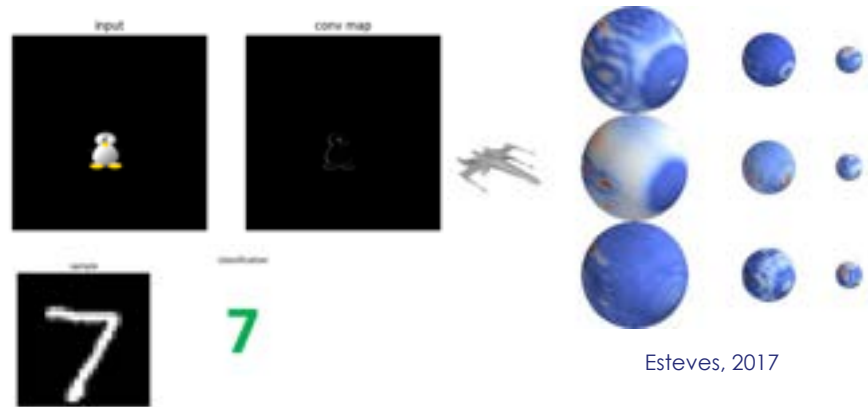
Geometric Deep Learning & Robustness

Geometric Deep Learning generalizes CNN

- **Equivariance** to other transforms than translations
- Operates in the **native geometry** of the data

Robustness is empirically improved

- Native **geometrical robustness** is achieved **by-design**
- Studies show that **local robustness** is also improved

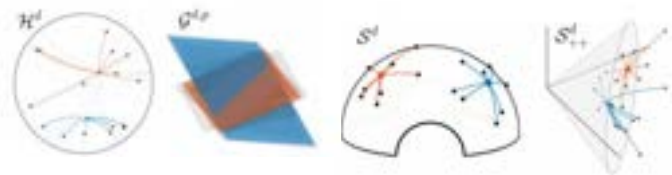


A rigorous analysis is however lacking

- Need for proper **metrics** to be defined
- **Stochastic framework** to be specified on **native spaces** of the data

Analysis outputs could help further increase robustness

- Refine the choice of the **underlying equivariance**
- Use **adversarial training** for a well chosen **distribution of perturbations**

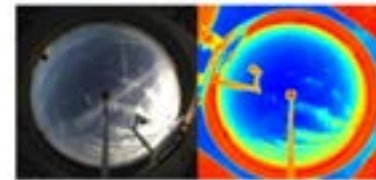


Applications for Sensors

Image processing

- Conventional + **Fish-Eye** images (patent pending)
- Investigate **Hybrid** approaches (AAAI-MLPS21)
 - LAS/TRT collaboration(IRS'21, GSI'21)

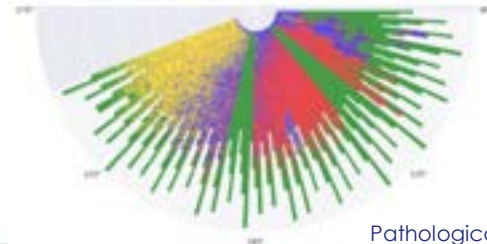
View from **FishEye Camera**



Contrails observation with skyimagers

Temporal signal processing

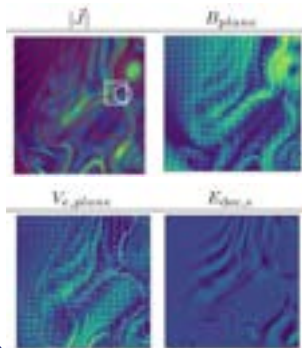
- **Radar** signal processing
 - LAS/TRT collaboration(IRS'21, GSI'21)



Pathological **radar clutter** classification,
Cabanes et Al, 2019

Equations de Maxwell-Vlasov

$$\begin{aligned} \frac{\partial f_e}{\partial t} + \mathbf{v}_e \cdot \nabla f_e - e \left(\mathbf{E} + \frac{\mathbf{v}_e \times \mathbf{B}}{c} \right) \cdot \frac{\partial f_e}{\partial \mathbf{p}} &= 0 \\ \frac{\partial f_i}{\partial t} + \mathbf{v}_i \cdot \nabla f_i + Ze \left(\mathbf{E} + \frac{\mathbf{v}_i \times \mathbf{B}}{c} \right) \cdot \frac{\partial f_i}{\partial \mathbf{p}} &= 0 \\ \nabla \times \mathbf{B} &= \frac{4\pi e}{c} \mathbf{j} + \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} \\ \nabla \times \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} \\ \nabla \cdot \mathbf{E} &= 4\pi \rho \\ \nabla \cdot \mathbf{B} &= 0 \end{aligned}$$



Hu et al., 2020

Simulation

- Solving **Partial Differential Equations** (LRASC internship)
- Also applicable to **Hamiltonian** systems

OPEN

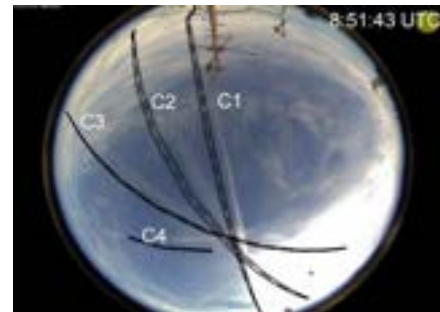
Toward Deep Learning for Contrails Detection by Fisheye Camera

Several possible applications of DL for contrails study

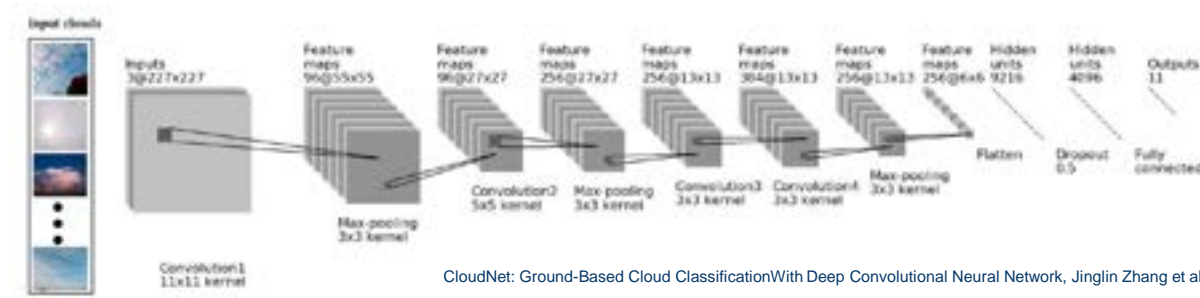
- Identification (vs. regular clouds)
- Tracking (originating planes)
- Properties analysis (altitudes, speeds etc.)
- Classification depending on radiative impact

Use of tailored Deep Learning algorithms

- Operating in **Fish-Eye geometry** for image processing
- Embed corresponding **equivariant layers**
- Couple with **physical models** and **meteorological data** where appropriate



Contrail study with ground-based Cameras, U. Schumann et al.



CloudNet: Ground-Based Cloud Classification With Deep Convolutional Neural Network, Jinglin Zhang et al.

$SU(1, 1)/U(1)$ Equivariance in Poincaré Disk

Fish-eye lens have hemispherical view

- Several projection methods exist, with corresponding distortion

Introduce a stereographic projection onto the Poincaré disk

- Project the hemispherical view onto the Poincaré disk \mathbb{D}^2 from the south pole
- Images can be represented as a signal $f: \mathbb{D}^2 \rightarrow \mathbb{R}^3$

Build $SU(1, 1)/U(1)$ equivariant convolutional layers

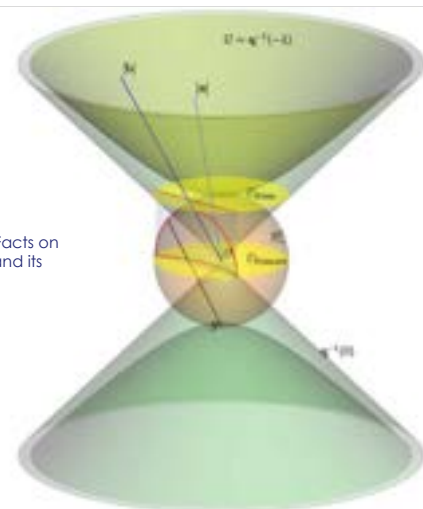
- Non-compact group ☹
- Synergy with the analytics being developed for radar Doppler signal processing

$$SU(1,1) = \left\{ g_{\alpha,\beta} = \begin{bmatrix} \alpha & \bar{\beta} \\ \beta & \bar{\alpha} \end{bmatrix}, |\alpha|^2 - |\beta|^2 = 1, \alpha, \beta \in \mathbb{C} \right\}$$

$$U(1) = \left\{ \begin{bmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{bmatrix}, |\alpha|^2 = 1, \alpha, \beta \in \mathbb{C} \right\}$$

$SU(1,1)$ acts on the Poincaré Disk by Mobius transform

$$g_{\alpha,\beta} \circ z = \frac{\alpha z + \beta}{\bar{\beta} z + \bar{\alpha}}$$



M. C. Escher, Circle Limit III



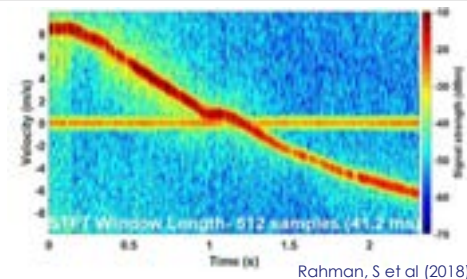
Motivations for Doppler Radar Processing

ML techniques have successfully been used for Doppler signal processing

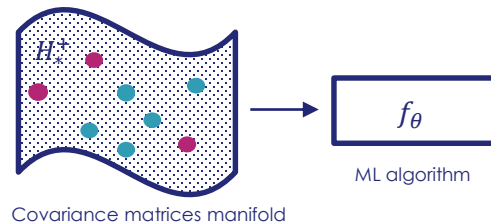
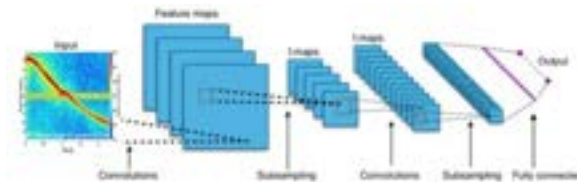
- Classical algorithms such as **CNN** and LSTM operating on **Doppler signatures**
- **Geometric Learning** operating on **complex covariance matrices**
 - D. Brooks et al, 2019
 - Y. Cabanes et al, 2019

Target here robustness to noise (e.g., thermal noise, clutter, etc.)

- Robustness is a **challenge for Deep Learning** (e.g., adversarial attacks)
- **CNN** are **locally robust to translation** of 2D images
- Want to **generalize geometric robustness** to algorithms operating on complex covariance matrices.



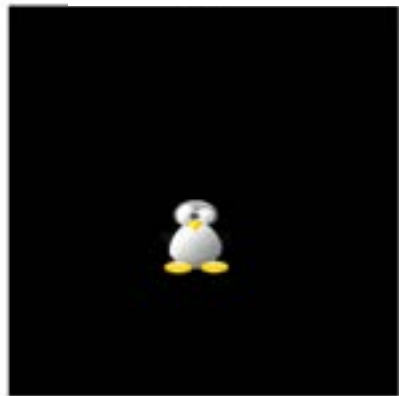
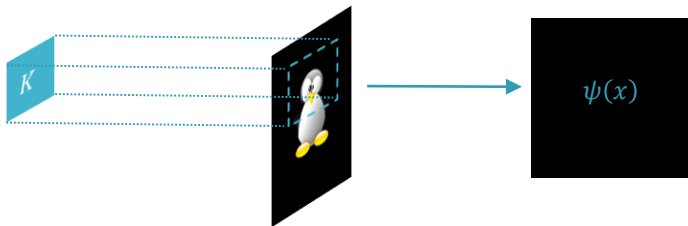
Rahman, S et al (2018)



Robustness property: $f_\theta(\Gamma_0) = f_\theta(\Gamma_1) \quad \forall \Gamma_0, \Gamma_1 \in H_*^+$,
with Γ_0 and Γ_1 "close enough"

2D-Convolution and Translation Equivariance

$$\forall x \in \mathbb{Z}^2, \quad \psi(x) = (I \star K)(x) = \sum_{y \in \mathbb{Z}^2} K(x - y)I(y)$$



Original image



Convolution feature map
with Sobel filter

$K: \mathbb{Z}^2 \rightarrow \mathbb{R}$ convolution kernel

$I: \mathbb{Z}^2 \rightarrow \mathbb{R}$ is the functional representation of the image

L_h : translation operator for $h \in \mathbb{Z}^2$, so that $L_h(f)(x) = f(x - h)$

Feature maps built from convolution kernels transform consistently with the original image translation

$$L_h \psi(x) = (L_h I \star K)(x)$$

CNN success for image processing tasks comes from this equivariance property, combined with the use of local filters, weights sharing, and pooling steps.

Convolution definition can also be seen as a discretized integral

$$\psi(x) = (I \star K)(x) = \int_{\mathbb{R}^2} K(x - y)I(y)dx dy$$

→ We want to generalize to other transforms than translations

Group Actions

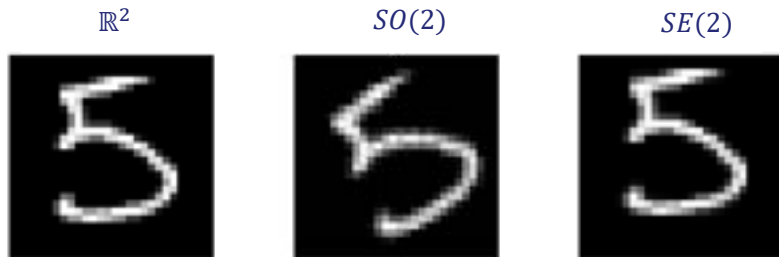
Group theory provides the mathematical tools to deal with structured transforms

- Transformations are represented by elements g of a group G
- Transformations $g \in G$ act on the data X as $g \circ X$

Examples of Groups of 2D-plane transforms

- Translation group \mathbb{R}^2 : $T_x \oplus_{\mathbb{R}^2} T_y = T_{x+y}$, for $x, y \in \mathbb{R}^2$
- Rotation group $SO(2)$: $R_{\theta_1} \oplus_{SO(2)} R_{\theta_2} = R_{\theta_{1,2}}$, for $\theta_1, \theta_2 \in [0, 2\pi]$ and $\theta_{1,2} = \theta_1 + \theta_2 \mod 2\pi$
- Special Euclidean group $SE(2)$: $H_{x,\theta_1} \oplus_{SE(2)} H_{y,\theta_2} = H_{R_{\theta_1}y+x, R_{\theta_{1,2}}}$ ($SE(2) = \mathbb{R}^2 \ltimes SO(2)$)

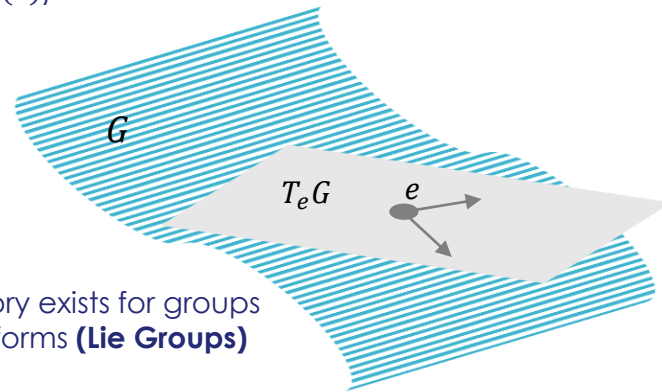
Examples of actions on a single image



Evariste Galois 1811 -1832



Sophus Lie 1842 -1899



A very powerful theory exists for groups of continuous transforms (**Lie Groups**)

Generalized Convolution

Convolution operator can be generalized to generic groups of transforms

$$\forall g \in G, \quad \psi_G(g) = (f \star_G K)(g) = \int_G K(h^{-1}g)f(h)d\mu^G(h)$$

$K: G \rightarrow \mathbb{R}$ convolution kernel

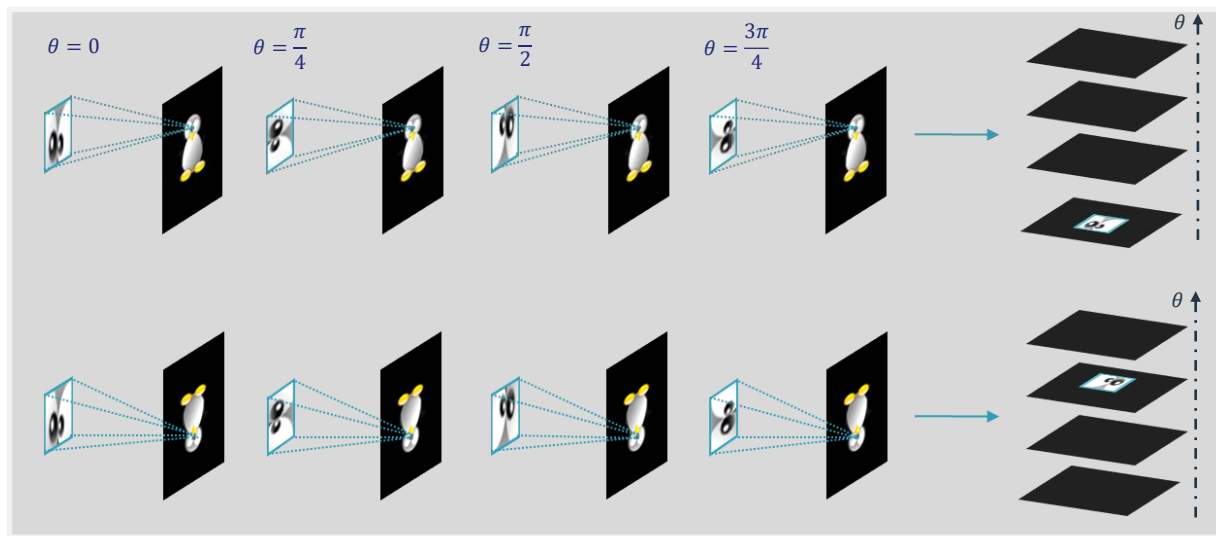
$f: G \rightarrow \mathbb{R}$ input feature map

L_h : translation operator for $h \in g$, so that $L_h(f)(g) = f(h^{-1}g)$

μ^G : Haar measure of the group G

Feature maps built from generalized convolutions are equivariant to the corresponding group action

$$L_h \psi_G(g) = (L_h f \star_G K)(g)$$



Equivariance of an $SE(2)$ convolution operator

- Kernel is rotated and 2D convolution is performed as usual
- A stack of 2D feature maps is obtained, indexed by the rotation angle
- Rotation of the input translates into a shift in the stack of feature maps

Motivations

The deployment of a radar could be alleviated by an automated recognition of pathological clutter

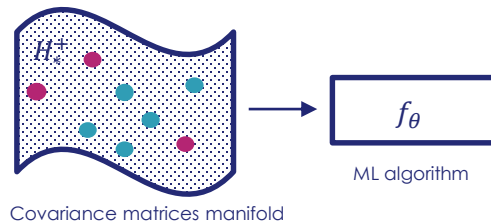
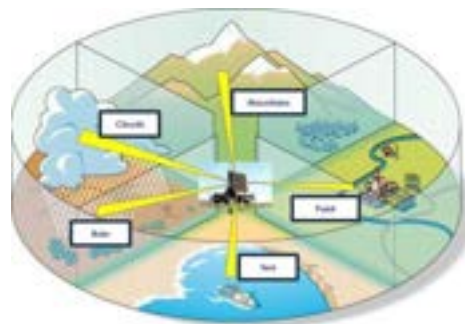
- Build a **Machine Learning** (ML) algorithm to identify specific **clutter characteristics** from their **Doppler spectrum** fluctuation

ML techniques have successfully been used for Doppler signal processing

- Classical algorithms such as **CNN** and **LSTM** operating on **Doppler signatures**
- **Geometric Learning** operating on **complex covariance matrices**
 - D. Brooks et al, 2019
 - Y. Cabanes et al, 2019

Target here robustness to noise (e.g., thermal noise)

- Robustness is a **challenge for Deep Learning** (e.g., adversarial attacks)
- **CNN** are **locally robust to translation** of 2D images
- Want to **generalize geometric robustness** to algorithms operating on complex covariance matrices.



Robustness property: $f_\theta(\Gamma_0) = f_\theta(\Gamma_1) \quad \forall \Gamma_0, \Gamma_1 \in H_*^+$, with Γ_0 and Γ_1 "close enough"

Radar Doppler Signal Classification

Represent signals by covariance matrices

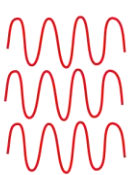
$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \dots & x_{1,p} \\ x_{2,1} & x_{2,2} & x_{2,3} & \dots & x_{2,p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & x_{n,3} & \dots & x_{n,p} \end{bmatrix}$$

Temporal axis
n = number of pulses
p = 10

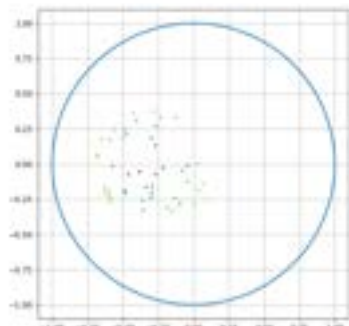
Distance axis
p = number of cells
 $x_{ij} \in \mathbb{C}$



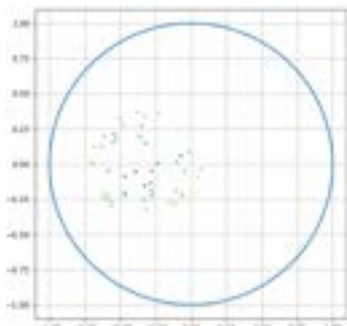
Pulses emitted



Build $SU(1,1)$ Equivariant NN

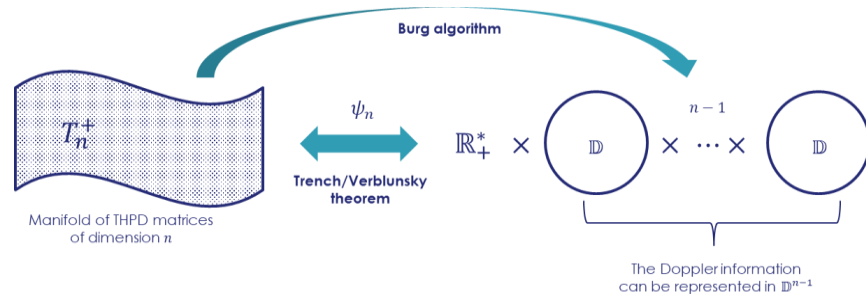


$SU(1,1)$
action
→

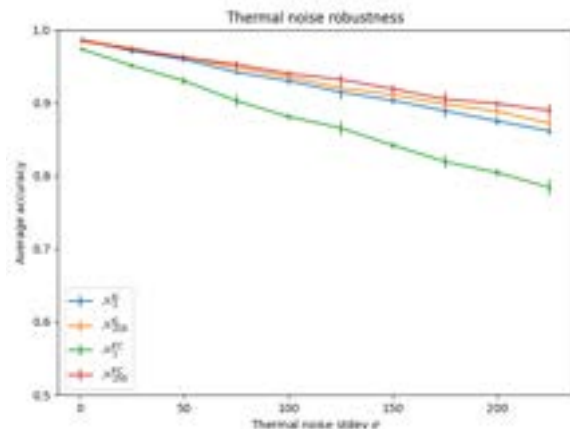
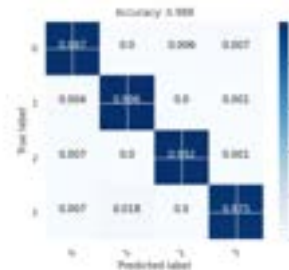


$SU(1,1)$ is **not compact** & \mathbb{D} is **not Euclidean** → **challenging set-up**

Hyperbolic embedding in a product space of Poincaré Disks

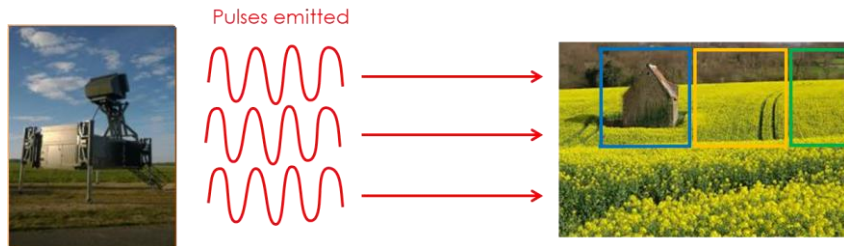


Achieve increased robustness



Anchor in Cabanes et al., 2019

- Process the data **cell-by-cell**
- Signal considered to be **a stationary, zero mean, autoregressive Gaussian process**



$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & x_{2,3} & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & x_{n,3} & \cdots & x_{n,p} \end{bmatrix}$$

Distance axis
 $p = \text{number of cells}$

Temporal axis
 $n = \text{number of pulses}$
 $n \approx 10$

$x_{i,j} \in \mathbb{C}$

Consider the autocorrelation matrix of each cell

- Stationarity assumption makes $R_{i,n}$ **Toeplitz Hermitian Positive Definite (THPD)**

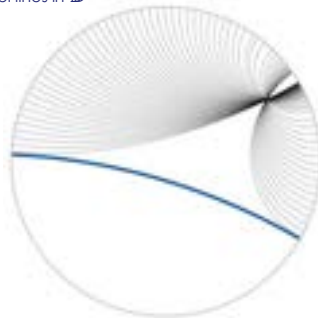
$$X_i = \begin{bmatrix} x_{1,i} \\ \vdots \\ x_{n,i} \end{bmatrix} \quad R_{i,n} = E[X_i X_i^+] = \begin{bmatrix} r_0 & r_1^* & \cdots & r_{n-1}^* \\ r_1 & r_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_1^* \\ r_{n-1} & \cdots & r_1 & r_0 \end{bmatrix} \quad \text{with } r_k = E[x_{m,i} x_{m-k,i}^*]$$

Several models of hyperbolic geometry exist

- Use the **Poincaré disk** model $\mathbb{D} = \{z \in \mathbb{C} / |z| < 1\}$
- Poincaré disk is endowed with a **non-Euclidean metric** $\rho_{\mathbb{D}}$

$$\rho_{\mathbb{D}}(z_1, z_2) = 2 \tanh^{-1} \left| \frac{z_2 - z_1}{1 - \bar{z}_1 z_2} \right|$$

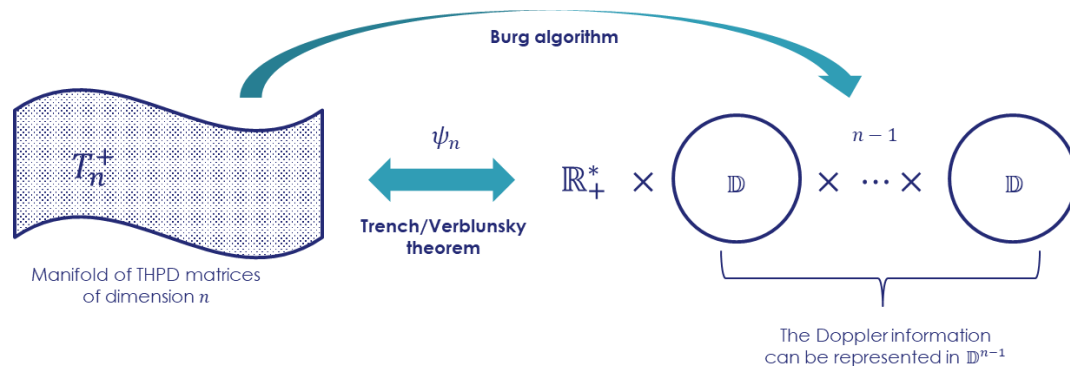
Parallel lines in \mathbb{D}



M. C. Escher, Circle Limit III



Embed THPD matrices into a product space of Poincaré disks



$$\psi_n: \begin{cases} T_n^+ \rightarrow \mathbb{R}_+^* \times \prod_{i=1}^{n-1} \mathbb{D} \\ \Gamma \rightarrow (p, \mu_1, \dots, \mu_{n-1}) \end{cases}$$

- Work with **rescaled autocorrelation matrices** and only focus on **reflection coefficients**

Group Action, Homogenous Space and Representation

Consider the Special Unitary Group (Lie group)

- $SU(1,1)$ is both a **group** and a **manifold**

$$SU(1,1) = \left\{ g_{\alpha,\beta} = \begin{bmatrix} \alpha & \bar{\beta} \\ \beta & \bar{\alpha} \end{bmatrix}, |\alpha|^2 - |\beta|^2 = 1, \alpha, \beta \in \mathbb{C} \right\}$$

Poincaré disk can be endowed with an action of $SU(1,1)$

- Acts **transitively** with \circ through **Mobius transforms** (isometries)
- Poincaré disk is an **homogenous space** $\mathbb{D} \simeq SU(1,1)/U(1)$

$$g_{\alpha,\beta} \circ z = \frac{\alpha z + \beta}{\bar{\beta} z + \bar{\alpha}} \quad U(1) = \left\{ \begin{bmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{bmatrix}, |\alpha|^2 = 1, \alpha, \beta \in \mathbb{C} \right\}$$

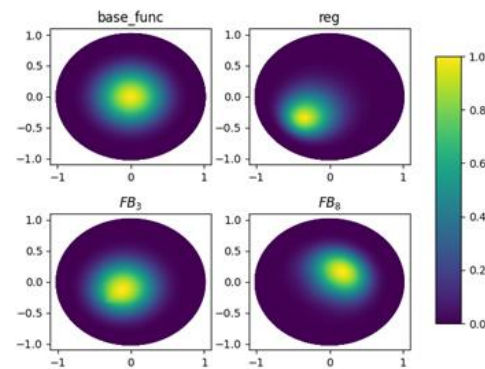
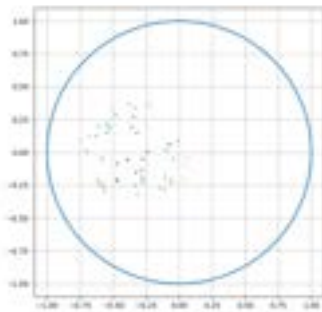
$SU(1,1)$ acts on Fock-Bargmann Hilbert spaces \mathcal{FB}_η

- Through the **UIR representation** ρ^η of $SU(1,1)$ on \mathcal{FB}_η

$$[\rho^\eta(g_{\alpha,\beta})(f)](z) = \frac{1}{(\alpha - \bar{\beta}z)^{2\eta}} f(g_{\alpha,\beta}^{-1} \circ z) \quad (\eta = 0 \rightarrow \text{regular rep.})$$

Considered THPD matrices can be seen as $n - 1$ coset elements

- Target local **robustness** with respect to $SU(1,1)$ action
- Use lifting to **use generalized convolution** within a **G-CNN architecture**



SU(1,1) Equivariance and G-CNN

Use of equivariant convolution operator on \mathbb{D}

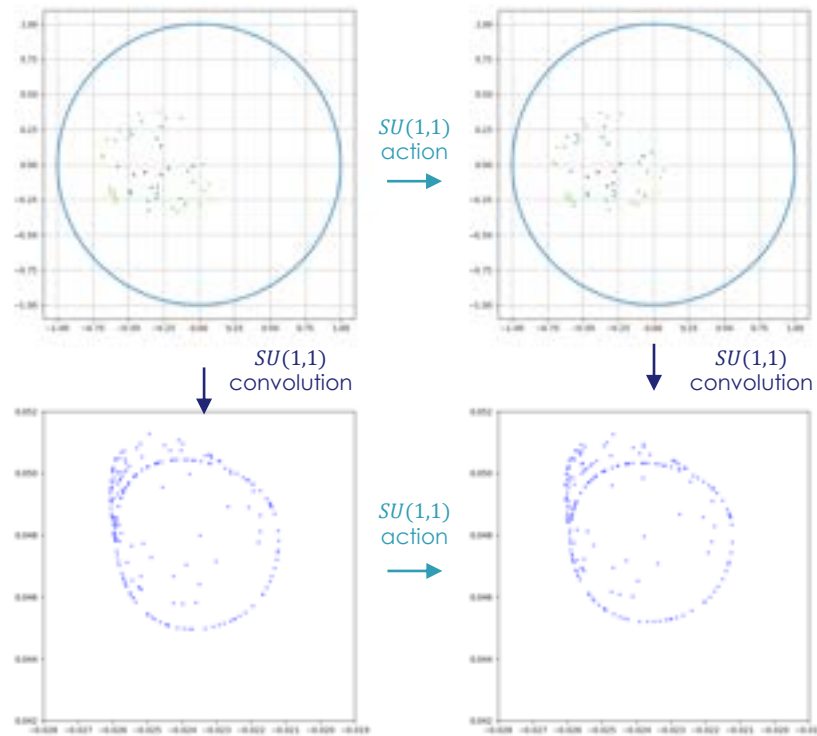
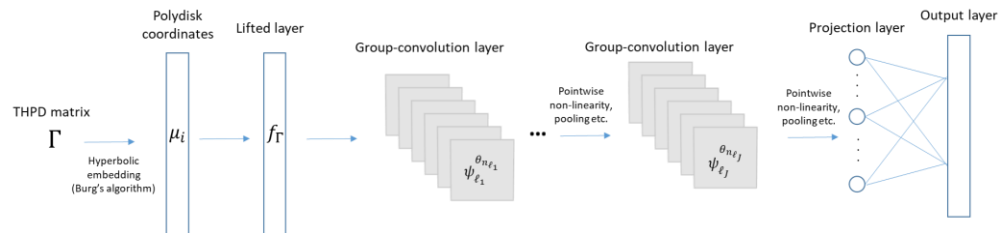
- \mathbb{D} is **not Euclidean** → use an adequate kernel
- Localization** allows **weights sharing** and integral definition

$$\forall z \in \mathbb{D}, \psi^\theta(z) = (f \star_{\mathbb{D}} k_\theta)(z) = \int_{h \in B_{\mathbb{D}}(z, M)} [\rho^\eta(h)(k_\theta)](z) [\rho^\eta(h)^{-1}(f)](0_{\mathbb{D}}) d\mu(h)$$

$$k_\theta(z) = \tilde{k}_\theta(\log_{\mathbb{D}}(z)) \quad \tilde{k}_\theta: \mathbb{R}^2 \rightarrow \mathbb{C} \quad \log_{\mathbb{D}}: \mathbb{D} \rightarrow \mathbb{R}^2 \text{ (Riemannian Logarithm)}$$

$$B_{\mathbb{D}}(z, M) = \{h \in SU(1,1) \mid \rho_{\mathbb{D}}(h \circ 0_{\mathbb{D}}, z) \leq M\} \quad \rho^\eta: \text{representation of } SU(1,1) \text{ on } \mathcal{FB}_\eta$$

ψ^θ operator are then used to build convolution layers



Equivariance of SU(1,1) convolution with a Gaussian Kernel

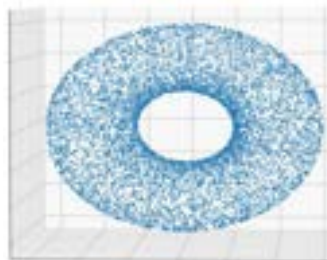
Numerical Considerations

Computing our convolution operator is generally challenging

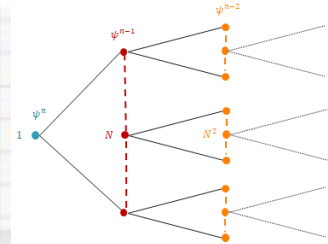
- $SU(1,1)$ is **not compact**

Can consider a Monte-Carlo approach

- Estimate with $\psi_N^\theta(z) = \frac{1}{N} \sum_{i=1}^N [\rho^\eta(g_i)(k)](z) [\rho^\eta(g_i)^{-1}(f)](0_{\mathbb{D}})$, where $g_i \sim \mu^{B_{\mathbb{D}}}(z, M)$
- Can use a **torus** parameterization (Poincaré disks as sections)
- But $\psi^{n+1}(z) = \Sigma (\psi^n \star_{\mathbb{D}} K)(z) \rightarrow$ **evaluation grids are dependent across layers**
- Exponential memory complexity with respect to depth



Sampling with Cartan parameterization of $SU(1,1)$

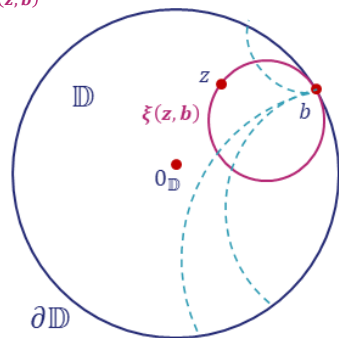


Size of the evaluation grids across layers

Can leverage on Helgason-Fourier (HF) Analysis

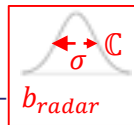
- HF transform is defined on $\mathbb{C} \times \partial\mathbb{D}$ for function $f: \mathbb{D} \rightarrow \mathbb{C}$ by $\hat{f}(\rho, b) = \int_{\mathbb{D}} f(z) e^{(1-i\rho)\langle z, b \rangle} dm(z)$ and an **inversion formula exists**
- **Convolution theorem** holds true (**regular rep.**): $\widehat{f \star_{\mathbb{D}} k}(\rho, b) = \hat{f}(\rho, b) \times \hat{k}(\rho, b)$
- 3 integrals computation but **can use a constant evaluation grid** across layers \rightarrow polynomial complexity
- **Expand kernel functions** on a functional basis (Bekkers, 2020) and **precompute HFT**.

Distance of $0_{\mathbb{D}}$ to the horocycle $\xi(z, b)$



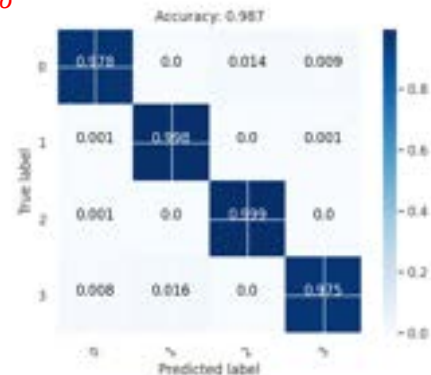
Experimental Proof-of-Concept

Provide a POC of our approach



Goal: assess robustness wrt. σ

- Work with **simulated data** ($n = 10$) $Z = \sqrt{\lambda} R^{1/2} x + b_{\text{radar}}$
- Consider **simple architectures** $\mathcal{N}_{\sigma}^{G,\eta}$ with one convolution layer with two filters modeled as a small MLP
- **Benchmark** with a **Fully Connected NN** $\mathcal{N}_{\sigma}^{FC}$ operating on the complex reflection coefficients (\approx same # of parameters)
- Average over 10 testing instances T_{σ} to **smooth-out the statistical noise**

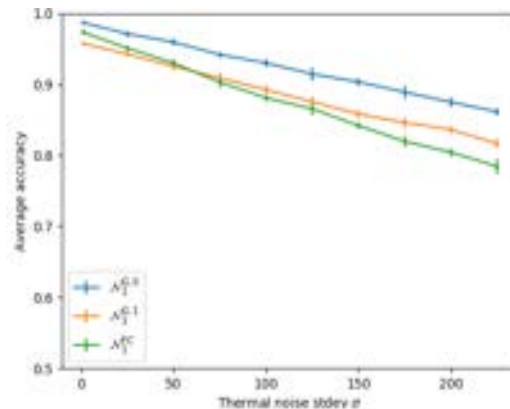


Confusion matrix of $\mathcal{N}_1^{G,0}$ on T_1

Consider training and evaluation with different values of σ

- $\mathcal{N}_1^{G,0}$ **significantly outperforms** \mathcal{N}_1^{FC} on T_{σ} as σ increases
- Robustness depends on the representation ρ^{η}
- **Calibrate** ρ^{η} on the targeted perturbation

→ Promising results from both accuracy and robustness standpoints



Evaluation of $\mathcal{N}_{\sigma}^{G,\eta}$ and $\mathcal{N}_{\sigma}^{FC}$ trained with several values of σ on several testing set T_{σ}

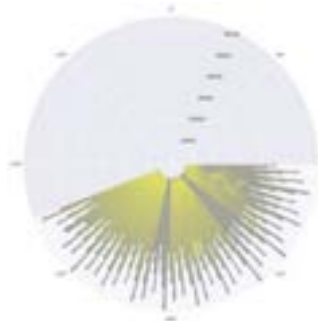
Radar Clutter Classification on Real Data

Preliminary work on real data recorded in Saint-Mandrier (FRANCE)

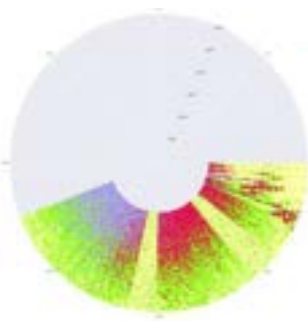
- Consider the labelling obtained through the clustering by Cabanes et al. (2019)



Ground map of Saint-Mandrier



Average power map

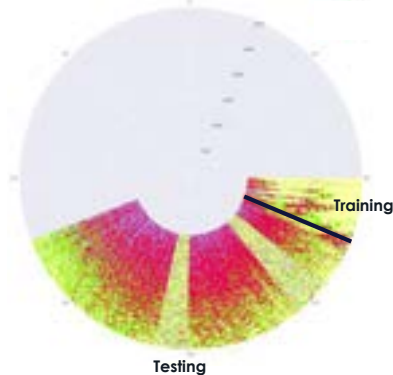
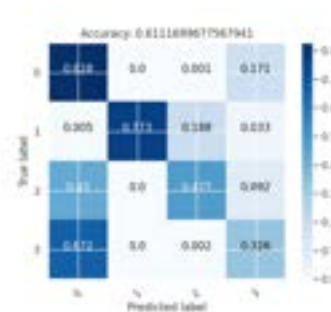


Clustering labels (Cabanes et al, 2019)

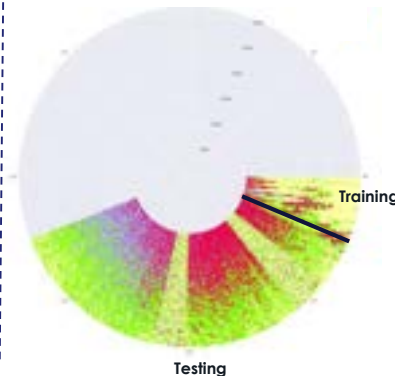
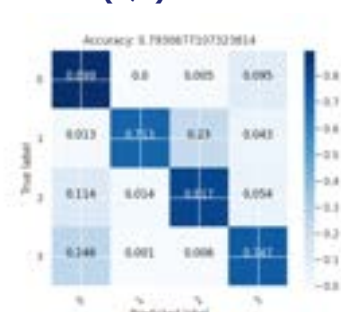
POC observations are confirmed in real world

- **Our SU(1,1)-CNN achieves 80% accuracy while a comparable FCNN only reaches 60%**

FCNN



SU(1,1)-CNN



Modern Sensor Processing based on Symplectic Model of Information

- Lie Group Based Equivariant GCNN for Adaptive Doppler Clutter Map
- Lie Group Based Frenet-Serret IEKF (Invariant Extended Kalman Filter) for tracking hyper-maneuvering targets
- Lie Group Based Target Recognition on Kinematics for Drone/Birds Classification
- Souriau Symplectic Model of Information for Lie Group Statistics and Machine Learning
 - Entropy as Casimir Function in Coadjoint Representation
 - Koszul-Fisher Metric on Lie Group
 - Covariant Maximum Entropy Density (Gauss Density) on Lie Group
 - Lie Groups Machine Learning

Kalman Filter on SE(3) Lie Group (Thales/Mines ParisTech Pilté PhD): IEFSKF: Invariant Extended Frenet-Serret Kalman Filter (1/3)

The state based on Frenet-Serret Model

- the state is $X_t = (R_t, x_t, \gamma_t, \tau_t, u_t)$

The kinematic Model

- The kinematic model is based on the Frenet-Serret frame evolution, and on the fact that the target is not allowed to slide during turns:

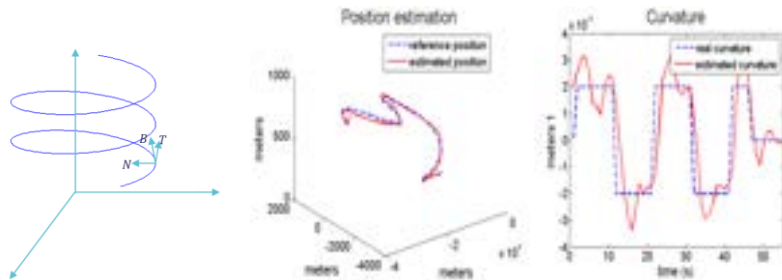
$$\frac{dx_t}{dt} = R_t(v_t + w_t^x), \frac{dR_t}{dt} = R_t(\omega_t + w_t^\omega)_\times, \frac{d\gamma_t}{dt} = 0 + w_t^\gamma, \frac{d\tau_t}{dt} = 0 + w_t^\tau, \frac{du_t}{dt} = 0 + w_t^u$$

With $v_t = [u_t, 0, 0]^T$, $\omega_t = [\tau_t, 0, \gamma_t]^T$, $(a)_\times \in \mathbb{R}^{3 \times 3}$ is the skew symmetric matrix associated to $a \in \mathbb{R}^3$.

- We can put part of the state (the rotation and translation, dim 6) into a matricial form:

$$\chi_t = \begin{pmatrix} R_t & x_t \\ 0_{1,3} & 1 \end{pmatrix}, \mu_t = \begin{pmatrix} R_t(\omega_t)_\times & v_t \\ 0_{1,3} & 0 \end{pmatrix}, \quad \frac{d}{dt} \chi_t = \chi_t(\mu_t + w_t^\chi)$$

- keep the other part (dim 3) in vectorial form: $z_t = (\gamma_t, \tau_t, u_t)$, $\frac{d}{dt} z_t = 0 + w_t^z$
- The cartesian measurement equation is $Y_n = x_{t_n} + V_n (= \chi_{t_n} d + V_n)$



THALES
Building a future we can all trust

Kalman Filter on SE(3) Lie Group (Thales/Mines ParisTech Pilté PhD): IEFSKF: Invariant Extended Frenet-Serret Kalman Filter (2/3)

IEFSKL Filter

[1] M.Pilté & S.Bonnabel & F.Barbaresco, *An Innovative Nonlinear Filter for Radar Kinematic Estimation of Maneuvering Targets in 2D*, 2017

[2] A.Barrau & S.Bonnabel *The Invariant Extended Kalman Filter as a stable observer*, 2016

- Evolution equation: $\frac{d\chi_t}{dt} = \chi_t(\mu_t + w_t^\chi)$, $\frac{dz_t}{dt} = \mathbf{0} + w_t^z$
- The χ -part of the state belongs to a matrix Lie group ($SE(3)$) and the z -part of the state belongs to a vectorial space.

- The error $\eta_t : \eta_t^\chi = \chi_t^{-1} \hat{\chi}_t$ and $\eta_t^z = \hat{z}_t - z_t$, $\eta_t = \begin{pmatrix} \eta_t^R \\ \eta_t^\chi \\ \eta_t^y \\ \eta_t^\tau \\ \eta_t^u \end{pmatrix} = \begin{pmatrix} R_t^T \hat{R}_t \\ R_t^T (\hat{\chi}_t - x_t) \\ \hat{y}_t - y_t \\ \hat{\tau}_t - \tau_t \\ \hat{u}_t - u_t \end{pmatrix}$



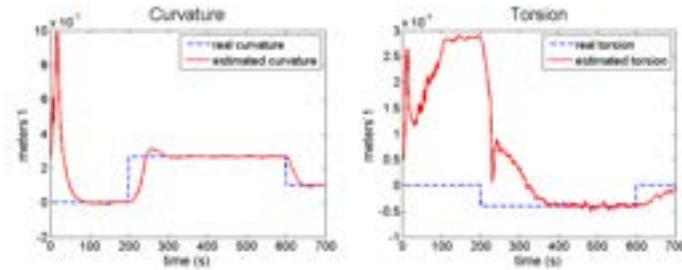
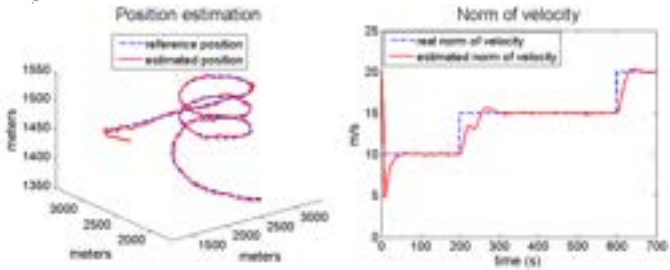
- The update step of the IEKF whenever a measurement Y_n is available and $t = t_n$:

$$\hat{\chi}_{t_n}^+ = \hat{\chi}_{t_n} \exp(L_n^\chi (\hat{\chi}_{t_n}^{-1} Y_n))$$

$$\hat{z}_{t_n}^+ = \hat{z}_{t_n} + L_n^z (\hat{\chi}_{t_n}^{-1} Y_n)$$

with L_n the Kalman gain

OPEN



Kalman Filter on SE(3) Lie Group (Thales/Mines ParisTech Pilté PhD): IEFSKF: Invariant Extended Frenet-Serret Kalman Filter (3/3)

- Linearized error ξ_t , such that $\eta_t^x \approx I + (\xi_t^x)_x$, and $\eta_t^z \approx \xi_t^z$, $\xi_t = \begin{pmatrix} \xi_t^R \\ \xi_t^x \\ \xi_t^y \\ \xi_t^z \\ \xi_t^u \end{pmatrix} \in \mathbb{R}^9$
- Linearized error evolution equation : $\frac{d\xi_t}{dt} = A_t \xi_t + w_t$ with A_t independent of \hat{R}_t, \hat{x}_t
- The Kalman gain L_n is computed by integrating the Riccati equation :

$$\frac{d}{dt} P_t = A_t P_t + P_t A_t^T + Q_t$$

$$S_n = H P_{t_n} H^T + \hat{R}_{t_n}^T N_n \hat{R}_{t_n}$$



$$L_n = P_{t_n} H^T S^{-1}$$

$$P_{t_n}^+ = (I_9 - L_n H) P_{t_n}$$

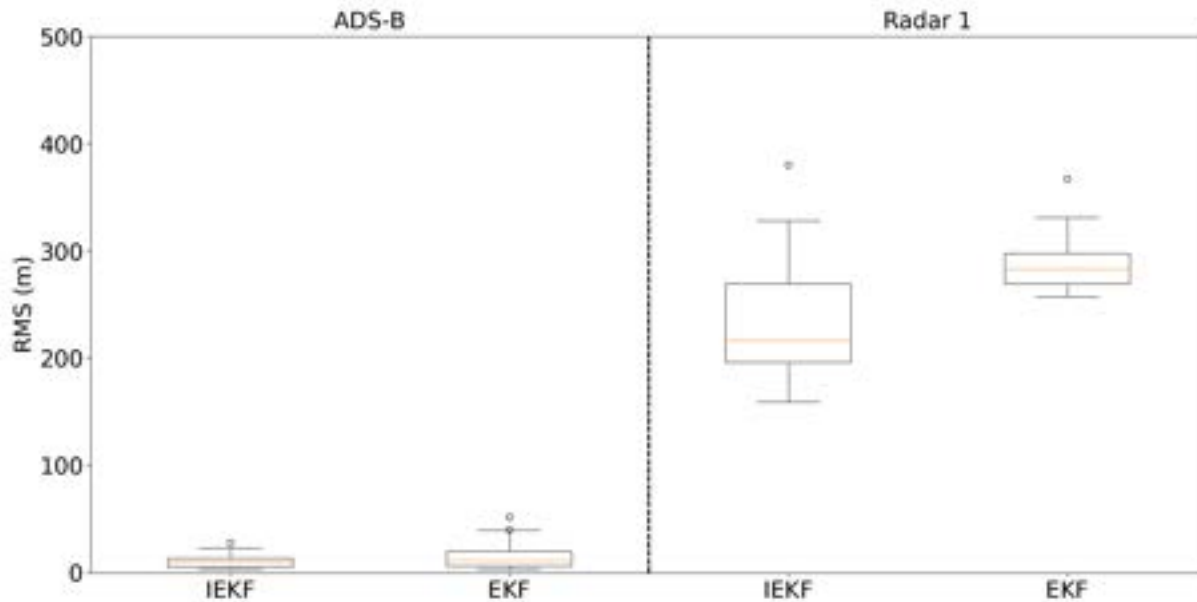
Q_t is the covariance of the process noise, and N_n is the covariance of the measurement noise, H is the measurement matrix.



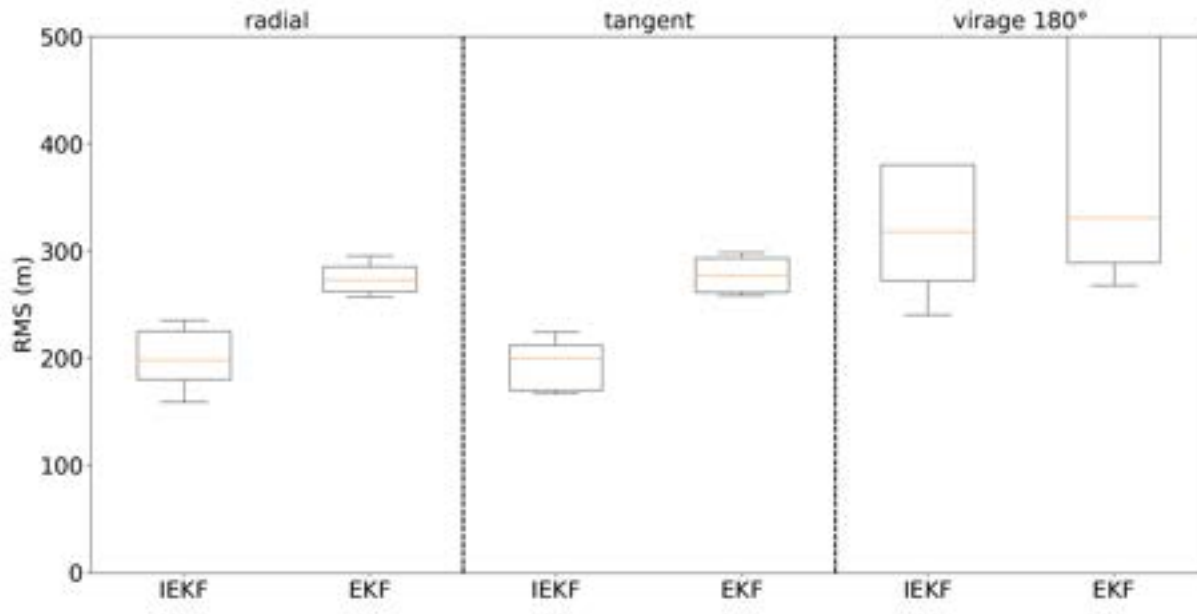
Lie Group Model (IEKF) versus Vector Space Model (EKF)

| | Linear Algebra (EKF) | Lie Algebra (IEKF)   |
|----------------------------|--|---|
| Kinematic Model | $\frac{d}{dt}X_t = f(X_t, w_t)$ | $\frac{d}{dt}\chi_t = \chi_t(v_t + w_t)$ |
| State Prediction | $\frac{d}{dt}\hat{X}_t = f(\hat{X}_t)$ | $\frac{d}{dt}\hat{\chi}_t = \hat{\chi}_t \hat{v}_t$ |
| Error definition | $\eta_t = \hat{X}_t - X_t$ | $\eta_t = \chi_t^{-1} \hat{\chi}_t$ |
| Error evolution | $\frac{d\eta_t}{dt} = f(\hat{X}_t) - f(X_t)$ | $\frac{d\eta_t}{dt} = \hat{v}_t \eta_t - \eta_t v_t - \eta_t w_t$ (autonome) |
| Linearized error evolution | $\frac{d\xi_t}{dt} = F_t(\hat{X}_t - X_t) + Q_t = F_t \xi_t + Q_t$ (F_t depend on predicted space) | $\frac{d\xi_t}{dt} = A_t \xi_t + Q_t$ (A_t independant of \hat{x}_t, \hat{R}_t) |
| Covariance definition | $P_t = Var(\xi_t)$ | $P_t = Var(\xi_t)$ |
| Covariance prediction | $\frac{dP_t}{dt} = F_t P_t + P_t F_t + Q_t$ | $\frac{dP_t}{dt} = A_t P_t + P_t A_t + Q_t$ |

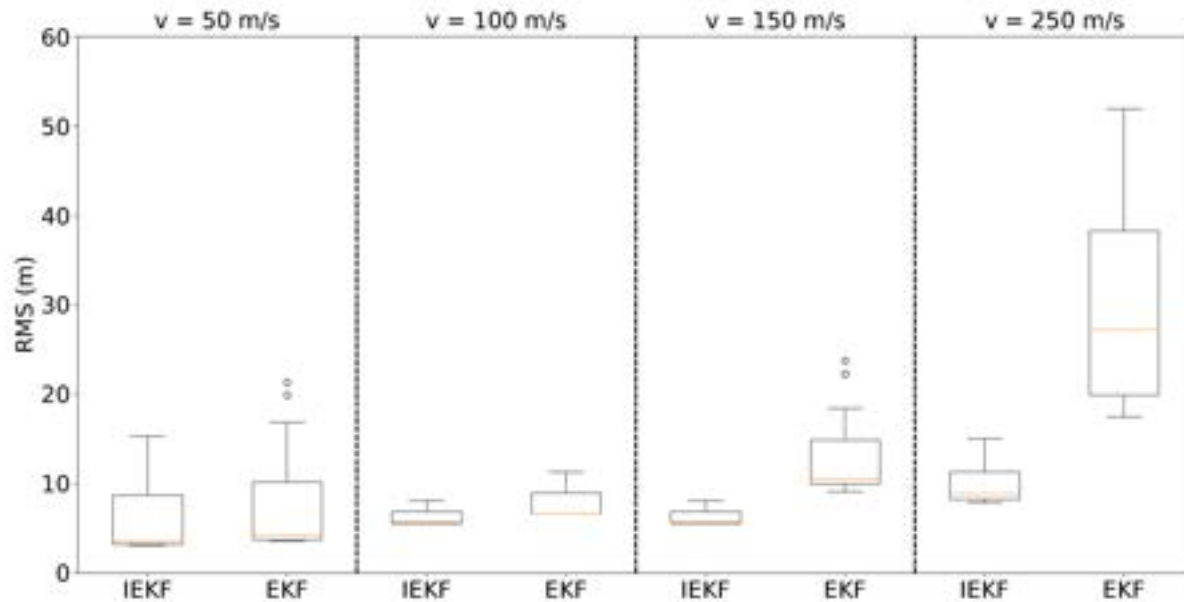
Lie Group Model (IEKF) versus Vector Space Model (EKF)



Lie Group Model (IEKF) versus Vector Space Model (EKF)

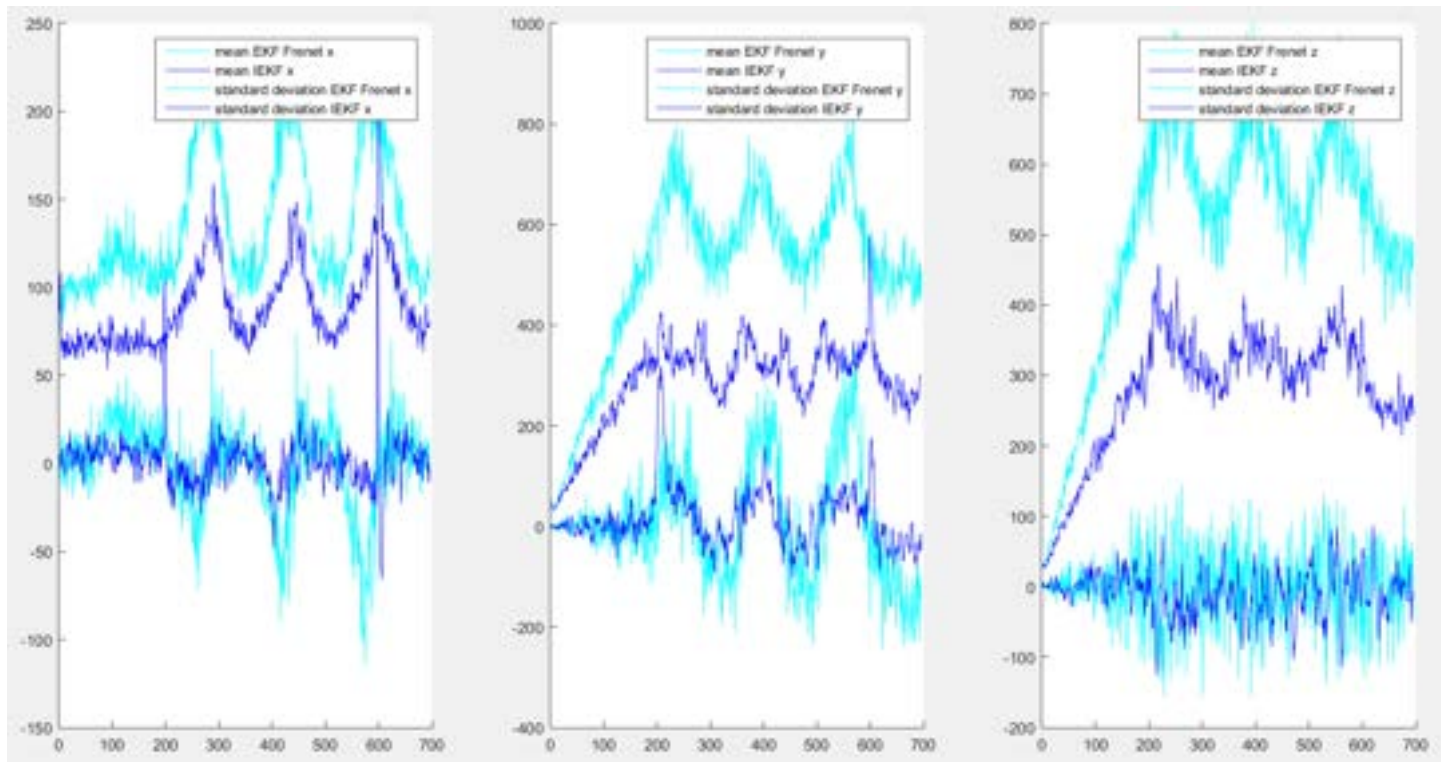


Lie Group Model (IEKF) versus Vector Space Model (EKF)



Lie Group Model (IEKF) versus Vector Space Model (EKF)

Superposition of mean and standard deviation



OPEN

Modern Sensor Processing based on Symplectic Model of Information

- Lie Group Based Equivariant GCNN for Adaptive Doppler Clutter Map
- Lie Group Based Frenet-Serret IEKF (Invariant Extended Kalman Filter) for tracking hyper-maneuvering targets
- Lie Group Based Target Recognition on Kinematics for Drone/Birds Classification
- Souriau Symplectic Model of Information for Lie Group Statistics and Machine Learning
 - Entropy as Casimir Function in Coadjoint Representation
 - Koszul-Fisher Metric on Lie Group
 - Covariant Maximum Entropy Density (Gauss Density) on Lie Group
 - Lie Groups Machine Learning

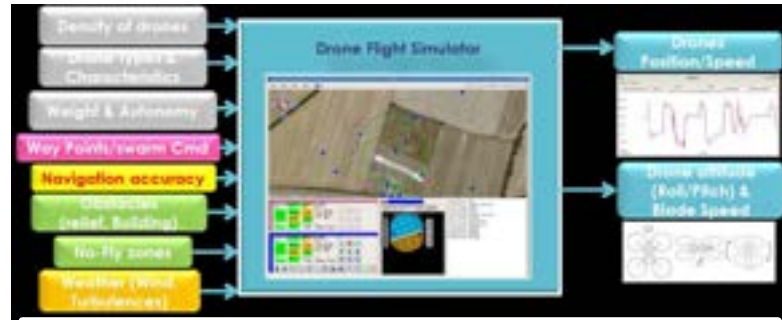
Drone Kinematics signature

- To improve the classification performance of drones when the Doppler signature of the blades is more difficult to characterize (blade shrouds, carbon blades, etc.), in addition to the Doppler signatures, the kinematic characteristics of the aircraft are considered.
- Tree boosting methods (XGBOOST type) are used by extracting statistical parameters on the time series of kinematic variables:
 - speed / acceleration / jerk on the 3 axes
 - horizontal velocity module
 - the 3D speed module
 - the horizontal 2D curvature
 - 3D curvature
 - the logarithm of the torsion of the 3D trajectory
- The classification is made on parameters of order statistics (median, quantile, L-moments, ...) estimated on the time series of these parameters.

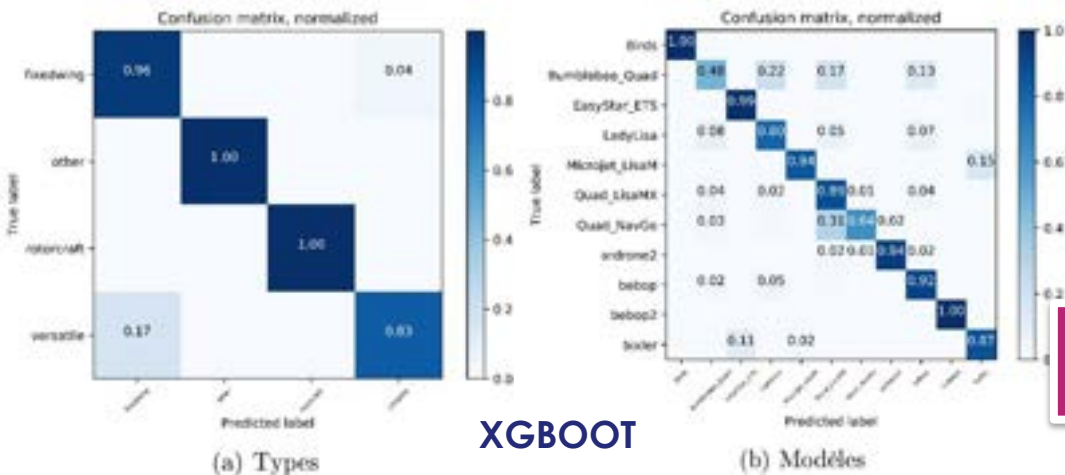
Target recognition on Kinematics by Gradient Boosting (XGBOOST)

Drone/Birds Recognition on Kinematics

- Drone trajectories and Kinematics are simulated by auto-pilot : **ENAC/TU Delft Paparazzi UAV**
- Birds trajectories and kinematics are characterized by GPS dataset on Birds: **MOVEBANK**
- Statistics features extraction (ordered statistics, L-moments, quantiles, ...) from time series of drone : speed / acceleration / jerk, 2D horizontal speed module, 3D speed module, 2D horizontal curvature, 3D curvature, torsion logarithm
- Python time series statistics: **lmoments**, **tsfresh**

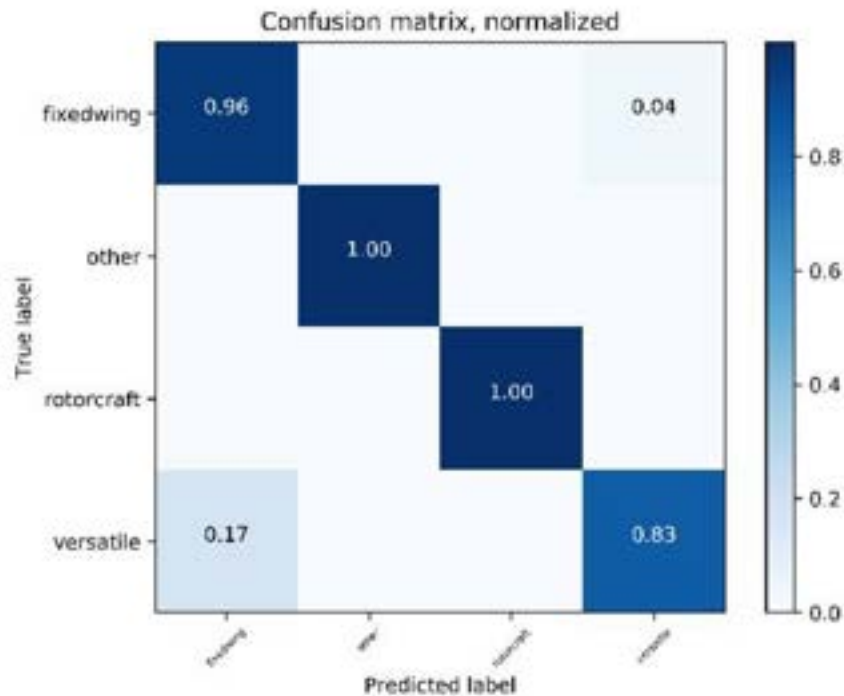


Drone Simulator:
https://wiki.paparazziuav.org/wiki/Main_Page

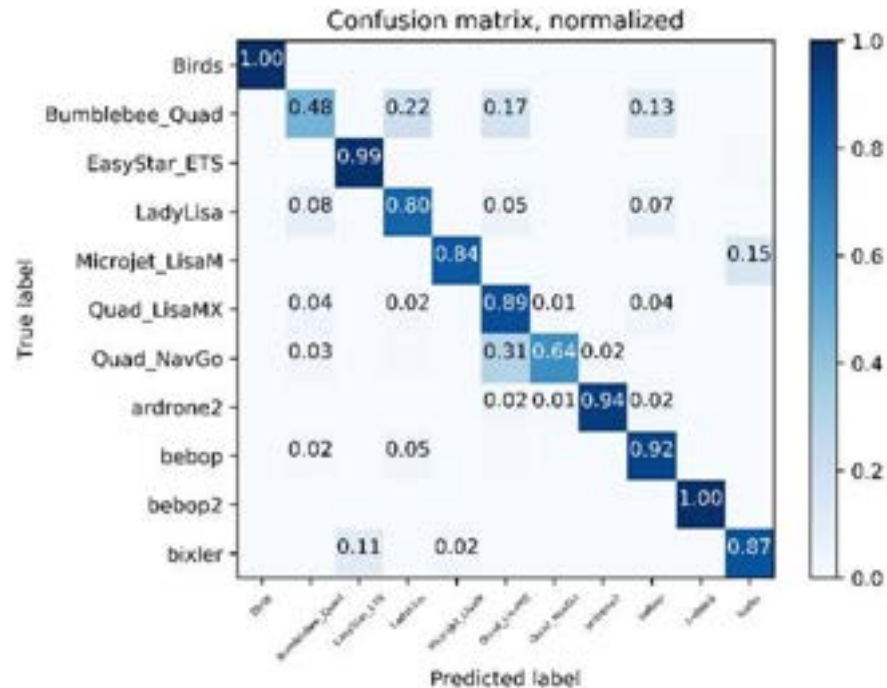


F. Barbaresco, D. Brooks, C. Adnet, Machine and Deep Learning for Drone Radar Recognition by Micro-Doppler and Kinematic criteria, IEEE Radar Conference, Florence, Sept.2020

Results for Pdetect=0.9

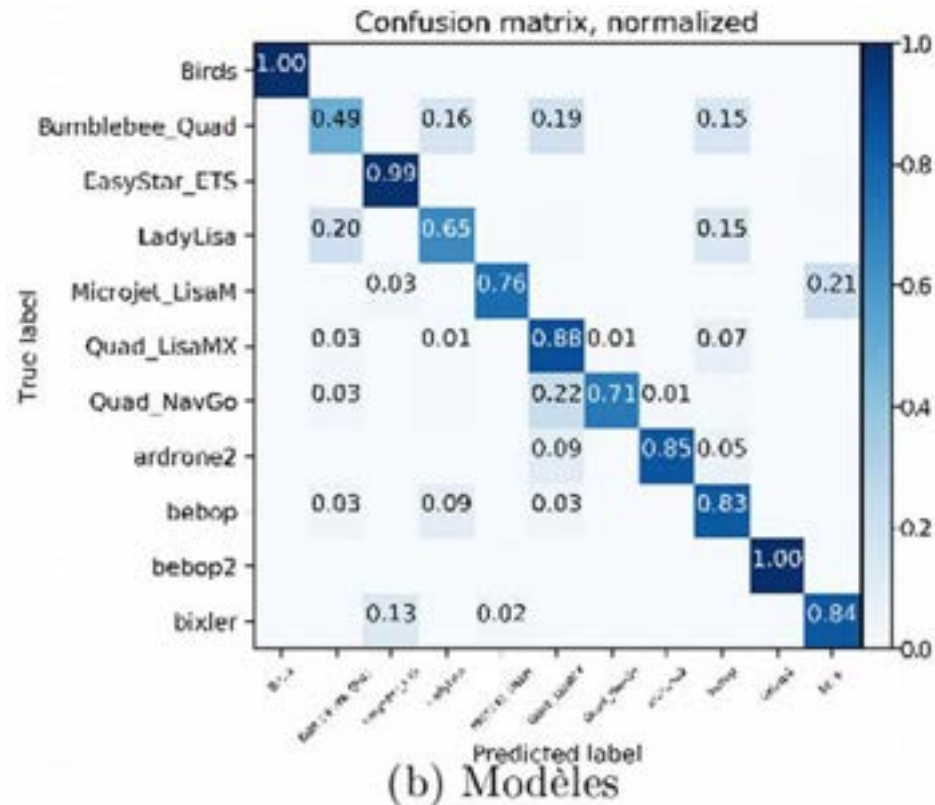
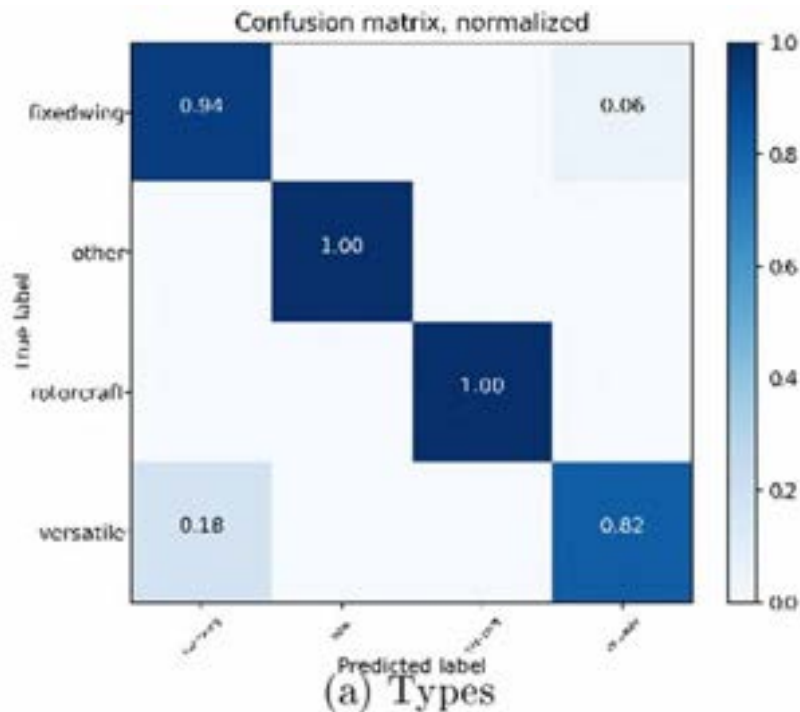


(a) Types



(b) Modèles

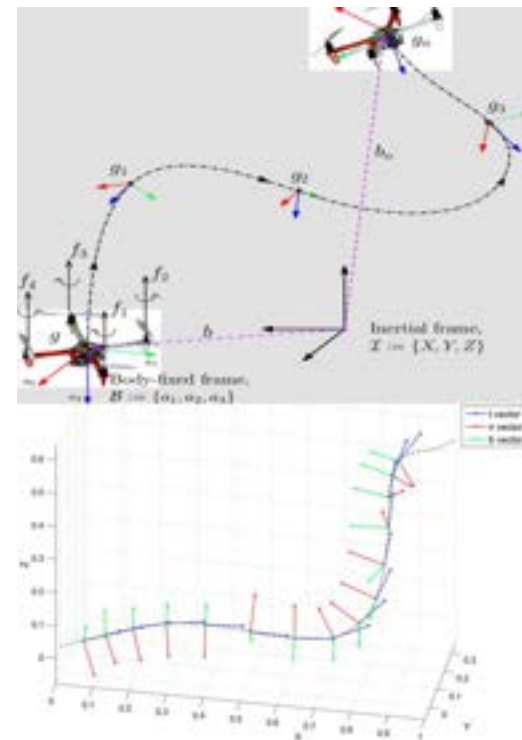
Results for Pdetect=0.6



3D trajectory and Frenet-Serret Frame

- When we consider a 3D trajectory of a mobile target, we can describe this curve by a time evolution of the local Frenet-Serret frame (local frame with tangent vector, normal vector and binormal vector). This frame evolution is described by the Frenet-Serret formula that gives the kinematic properties of the target moving along the continuous, differentiable curve in 3D Euclidean space \mathbb{R}^3 . More specifically, the formulas describe the derivatives of the so-called tangent, normal, and binormal unit vectors in terms of each other.

$$\frac{d}{dt} \begin{pmatrix} \vec{t} \\ \vec{n} \\ \vec{b} \end{pmatrix} = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \gamma \\ 0 & -\gamma & 0 \end{bmatrix} \begin{pmatrix} \vec{t} \\ \vec{n} \\ \vec{b} \end{pmatrix} \quad \text{with} \quad \begin{cases} \kappa : \text{curvature} \\ \gamma : \text{torsion} \end{cases}$$



3D trajectory curve

- we will consider motions determined by exponentials of paths in the Lie algebra. Such a motion is determined by a unit speed space-curve $\tau(t)$. Now in a Frenet-Serret motion a point in the moving body moves along the curve and the coordinate frame in the moving body remains aligned with the tangent \vec{t} , normal \vec{n} , and binormal \vec{b} , of the curve. Using the 4-dimensional representation of the Lie Group $SE(3)$, the motion can be specified as :

$$G(t) = \begin{pmatrix} R(t) & \tau(t) \\ 0 & 1 \end{pmatrix} \in SE(3)$$

- where $\tau(t)$ is the curve and the rotation matrix has the unit vectors \vec{t} , \vec{n} , and \vec{b} as columns:

$$R(t) = \begin{pmatrix} \vec{t} & \vec{n} & \vec{b} \end{pmatrix} \in SO(3)$$

Time evolution of Frenet-Serret Frame

- If we introduce the Darboux vector $\vec{\omega} = \gamma \vec{t} + \kappa \vec{b}$ that we can rewrite from Frenet-Serret Formulas :

$$\frac{d\vec{t}}{dt} = \vec{\omega} \times \vec{t} \quad , \quad \frac{d\vec{n}}{dt} = \vec{\omega} \times \vec{n} \quad , \quad \frac{d\vec{b}}{dt} = \vec{\omega} \times \vec{b}$$

- Then, we can write with Ω is the 3×3 anti-symmetric matrix corresponding to $\vec{\omega}$:

$$\frac{dR}{dt} = \Omega R$$

- We note that $\frac{d\tau(t)}{dt} = \vec{t}$ and $\frac{d\vec{\omega}}{dt} = \frac{d\gamma}{dt} \vec{t} + \frac{d\kappa}{dt} \vec{b}$

- The instantaneous twist of the motion $G(t)$ is given by:

$$S_d = \frac{dG(t)}{dt} G^{-1}(t) = \begin{pmatrix} \Omega & \nu \\ 0 & 0 \end{pmatrix}$$

Instantaneous twist

- This is the Lie algebra element corresponding to the tangent vector to the curve $G(t)$. It is well known that elements of the Lie algebra $se(3)$ can be described as lines with a pitch. The fixed axode of a motion $G(t) \in SE(3)$ is given by the axis of S_d as t varies. The instantaneous twist in the moving reference frame is given by

$S_b = G^{-1}(t)S_d G(t)$, that is, by the adjoint action on the twist in the fixed frame. The instantaneous twist S_b can also be found from the relation:

$$S_b = G^{-1}(t) \frac{dG(t)}{dt}$$

$$S_b = G^{-1} \frac{dG}{dt} = \begin{pmatrix} R^T & -R^T \tau \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Omega R & \vec{t} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} R^T \Omega R & R \vec{t} \\ 0 & 0 \end{pmatrix}$$

Matrix Lie Group SE(3) for Kinematic Data

Trajectory as a time series of Matrix SE(3) Lie groups

- We can observe that we could describe a 3D trajectory by a time series of SE(3) Lie group elements:

$$SE(3) = \left\{ \begin{bmatrix} R & \tau \\ 0 & 1 \end{bmatrix} / R \in SO(3), \tau \in R^3 \right\}$$

$$\text{with } SO(3) = \left\{ R / R^T R = R R^T = I, \det^2 R = 1 \right\}$$

- Then, the trajectory will be given by the following time series :

$$\left\{ \begin{bmatrix} R_1 & \tau_1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} R_2 & \tau_2 \\ 0 & 1 \end{bmatrix}, \dots, \begin{bmatrix} R_n & \tau_n \\ 0 & 1 \end{bmatrix} \right\} \in SE(3)^n$$

Modern Sensor Processing based on Symplectic Model of Information

- Lie Group Based Equivariant GCNN for Adaptive Doppler Clutter Map
- Lie Group Based Frenet-Serret IEKF (Invariant Extended Kalman Filter) for tracking hyper-maneuvering targets
- Lie Group Based Target Recognition on Kinematics for Drone/Birds Classification
- Souriau Symplectic Model of Information for Lie Group Statistics and Machine Learning
 - Entropy as Casimir Function in Coadjoint Representation
 - Koszul-Fisher Metric on Lie Group
 - Covariant Maximum Entropy Density (Gauss Density) on Lie Group
 - Lie Groups Machine Learning

Fisher Metric and Fréchet-Darmois (Cramer-Rao) Bound

- Cramer-Rao –Fréchet-Darmois Bound has been introduced by Fréchet in 1939 and by Rao in 1945 as inverse of the Fisher Information Matrix: $I(\theta)$

$$R_{\hat{\theta}} = E\left[(\theta - \hat{\theta})(\theta - \hat{\theta})^+\right] \geq I(\theta)^{-1} \quad [I(\theta)]_{i,j} = -E\left[\frac{\partial^2 \log p_{\theta}(z)}{\partial \theta_i \partial \theta_j^*}\right]$$

- Rao has proposed to introduced an invariant metric in parameter space of density of probabilities (axiomatised by N. Chentsov):

$$ds_{\theta}^2 = \text{Kullback_Divergence}(p_{\theta}(z), p_{\theta+d\theta}(z))$$

$$ds_{\theta}^2 = -\int p_{\theta}(z) \log \frac{p_{\theta+d\theta}(z)}{p_{\theta}(z)} dz$$

$$ds_{\theta}^2 \underset{\text{Taylor}}{\approx} \sum_{i,j} g_{ij} d\theta_i d\theta_j^* = \sum_{i,j} [I(\theta)]_{i,j} d\theta_i d\theta_j^* = d\theta^+ . I(\theta) . d\theta$$

$$w = W(\theta)$$

$$\Rightarrow ds_w^2 = ds_{\theta}^2$$

Distance Between Gaussian Density with Fisher Metric

Fisher Matrix for Gaussian Densities:

$$I(\theta) = \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{bmatrix} \quad \text{avec} \quad E\left[(\theta - \hat{\theta})(\theta - \hat{\theta})^T\right] \geq I(\theta)^{-1} \quad \text{et} \quad \theta = \begin{pmatrix} m \\ \sigma \end{pmatrix}$$

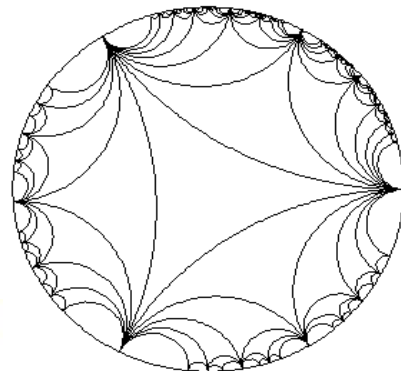
➤ Fisher matrix induced the following differential metric :

$$ds^2 = d\theta^T \cdot I(\theta) \cdot d\theta = \frac{dm^2}{\sigma^2} + 2 \cdot \frac{d\sigma^2}{\sigma^2} = \frac{2}{\sigma^2} \left[\left(\frac{dm}{\sqrt{2}} \right)^2 + (d\sigma)^2 \right]$$

➤ Poincaré Model of upper half-plane and unit disk

$$z = \frac{m}{\sqrt{2}} + i \cdot \sigma \quad \omega = \frac{z - i}{z + i} \quad (|\omega| < 1)$$

$$\Rightarrow ds^2 = 8 \cdot \frac{|d\omega|^2}{(1 - |\omega|^2)^2}$$



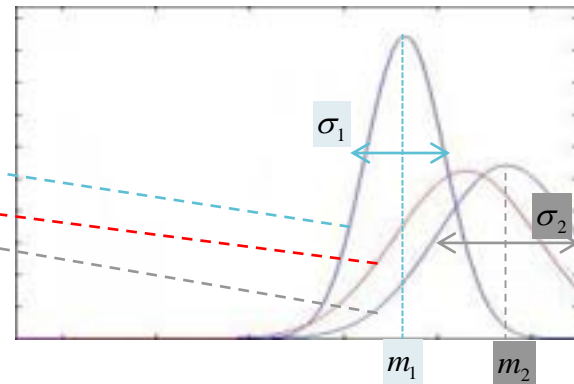
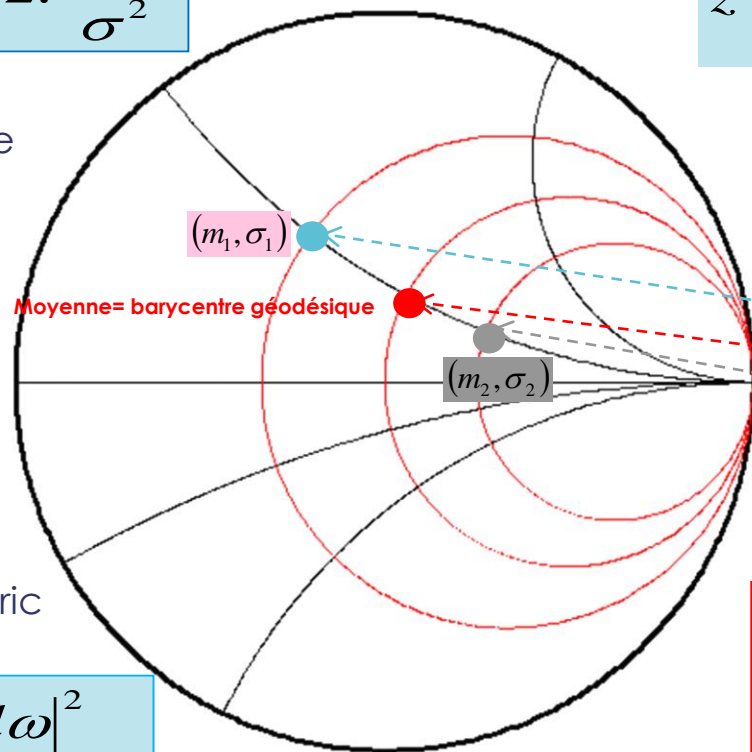
1 monivariate gaussian = 1 point in Poincaré unit disk

$$ds^2 = \frac{dm^2}{\sigma^2} + 2 \cdot \frac{d\sigma^2}{\sigma^2}$$

Fisher Metric in
Poincaré Half-Plane

$$z = \frac{m}{\sqrt{2}} + i \cdot \sigma$$

$$\omega = \frac{z - i}{z + i} \quad (|\omega| < 1)$$



Poincaré-Fisher metric
In Unit Disk

$$ds^2 = 8 \cdot \frac{|d\omega|^2}{(1 - |\omega|^2)^2}$$

$$d^2(\{m_1, \sigma_1\}, \{m_2, \sigma_2\}) = 2 \cdot \left(\log \frac{1 + \delta(\omega^{(1)}, \omega^{(2)})}{1 - \delta(\omega^{(1)}, \omega^{(2)})} \right)^2$$

with $\delta(\omega^{(1)}, \omega^{(2)}) = \left| \frac{\omega^{(1)} - \omega^{(2)}}{1 - \omega^{(1)} \omega^{(2)*}} \right|$

Gradient descent for Learning

- Information geometry has been derived from invariant geometrical structure involved in statistical inference. The Fisher metric defines a Riemannian metric as the Hessian of two dual potential functions, linked to dually coupled affine connections in a manifold of probability distributions. With the Souriau model, this structure is extended preserving the Legendre transform between two dual potential function parametrized in Lie algebra of the group acting transitively on the homogeneous manifold.
- Classically, to optimize the parameter θ of a probabilistic model, based on a sequence of observations y_t , is an online gradient descent with learning rate η_t , and the loss function $l_t = -\log p(y_t / \hat{y}_t)$:

$$\theta_t \leftarrow \theta_{t-1} - \eta_t \frac{\partial l_t(y_t)^T}{\partial \theta}$$

Information Geometry & Natural Gradient

- This simple gradient descent has a first drawback of using the same non-adaptive learning rate for all parameter components, and a second drawback of non invariance with respect to parameter re-encoding inducing different learning rates. **S.I. Amari** has introduced the **natural gradient** to preserve this invariance to be insensitive to the characteristic scale of each parameter direction. The gradient descent could be corrected by $I(\theta)^{-1}$ where I is the **Fisher information matrix** with respect to parameter θ , given by:

$$\theta_t \leftarrow \theta_{t-1} - \eta_t I(\theta_{t-1})^{-1} \frac{\partial l_t(y_t)^T}{\partial \theta}$$

$$I(\theta) = [g_{ij}]$$

$$\text{with } g_{ij} = \left[-E_{y \approx p(y/\theta)} \left[\frac{\partial^2 \log p(y/\theta)}{\partial \theta_i \partial \theta_j} \right] \right]_{ij} = \left[E_{y \approx p(y/\theta)} \left[\frac{\partial \log p(y/\theta)}{\partial \theta_i} \frac{\partial \log p(y/\theta)}{\partial \theta_j} \right] \right]_{ij}$$

Information Geometry & Machine Learning : Legendre structure

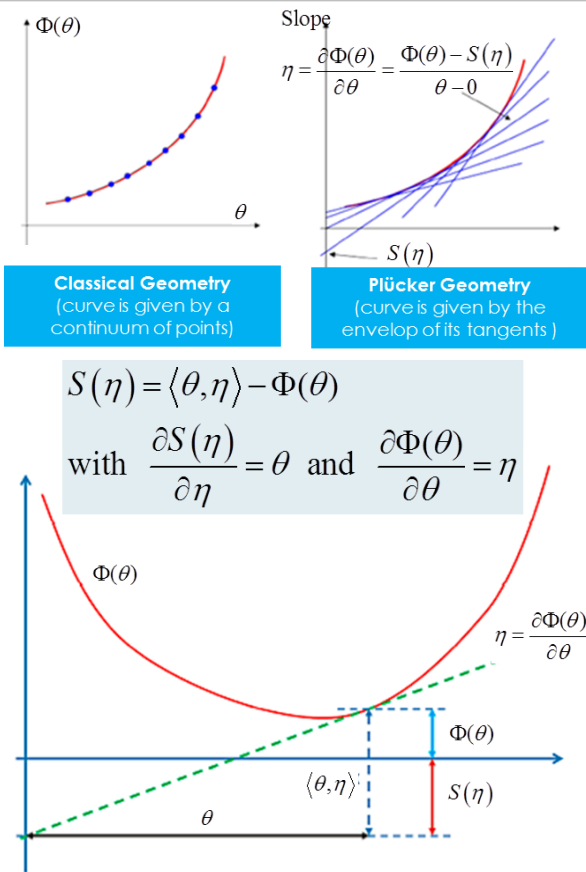
Legendre Transform, Dual Potentials & Fisher Metric

- **S.I. Amari** has proved that the Riemannian metric in an exponential family is the **Fisher information matrix** defined by:

$$g_{ij} = - \left[\frac{\partial^2 \Phi}{\partial \theta_i \partial \theta_j} \right]_{ij} \quad \text{with} \quad \Phi(\theta) = -\log \int_R e^{-\langle \theta, y \rangle} dy$$

- and the dual potential, the **Shannon entropy**, is given by the **Legendre transform**:

$$S(\eta) = \langle \theta, \eta \rangle - \Phi(\theta) \quad \text{with} \quad \eta_i = \frac{\partial \Phi(\theta)}{\partial \theta_i} \quad \text{and} \quad \theta_i = \frac{\partial S(\eta)}{\partial \eta_i}$$



Notation !

- I use notation as used by Koszul and Souriau which is not the most classical one

$$Ad_g^* = \left(Ad_{g^{-1}} \right)^*$$

with

$$\left\langle Ad_g^* F, Y \right\rangle = \left\langle F, Ad_{g^{-1}} Y \right\rangle, \forall g \in G, Y \in \mathfrak{g}, F \in \mathfrak{g}^*$$

Statistical Mechanics, Dual Potentials & Fisher Metric

- In geometric statistical mechanics, **J.M. Souriau** has developed a “**Lie groups thermodynamics**” of dynamical systems where the (maximum entropy) **Gibbs density is covariant** with respect to the action of the Lie group. In the Souriau model, previous structures of information geometry are preserved:

$$I(\beta) = -\frac{\partial^2 \Phi}{\partial \beta^2} \text{ with } \Phi(\beta) = -\log \int_M e^{-\langle \beta, U(\xi) \rangle} d\lambda \quad U: M \rightarrow \mathfrak{g}^*$$

$$S(Q) = \langle \beta, Q \rangle - \Phi(\beta) \text{ with } Q = \frac{\partial \Phi(\beta)}{\partial \beta} \in \mathfrak{g}^* \text{ and } \beta = \frac{\partial S(Q)}{\partial Q} \in \mathfrak{g}$$



Jean-Marie Souriau

- In the Souriau **Lie groups thermodynamics** model, β is a “geometric” (Planck) temperature, element of Lie algebra \mathfrak{g} of the group, and Q is a “geometric” heat, element of dual Lie algebra \mathfrak{g}^* of the group.

Statistical Mechanics & Invariant Souriau-Fisher Metric

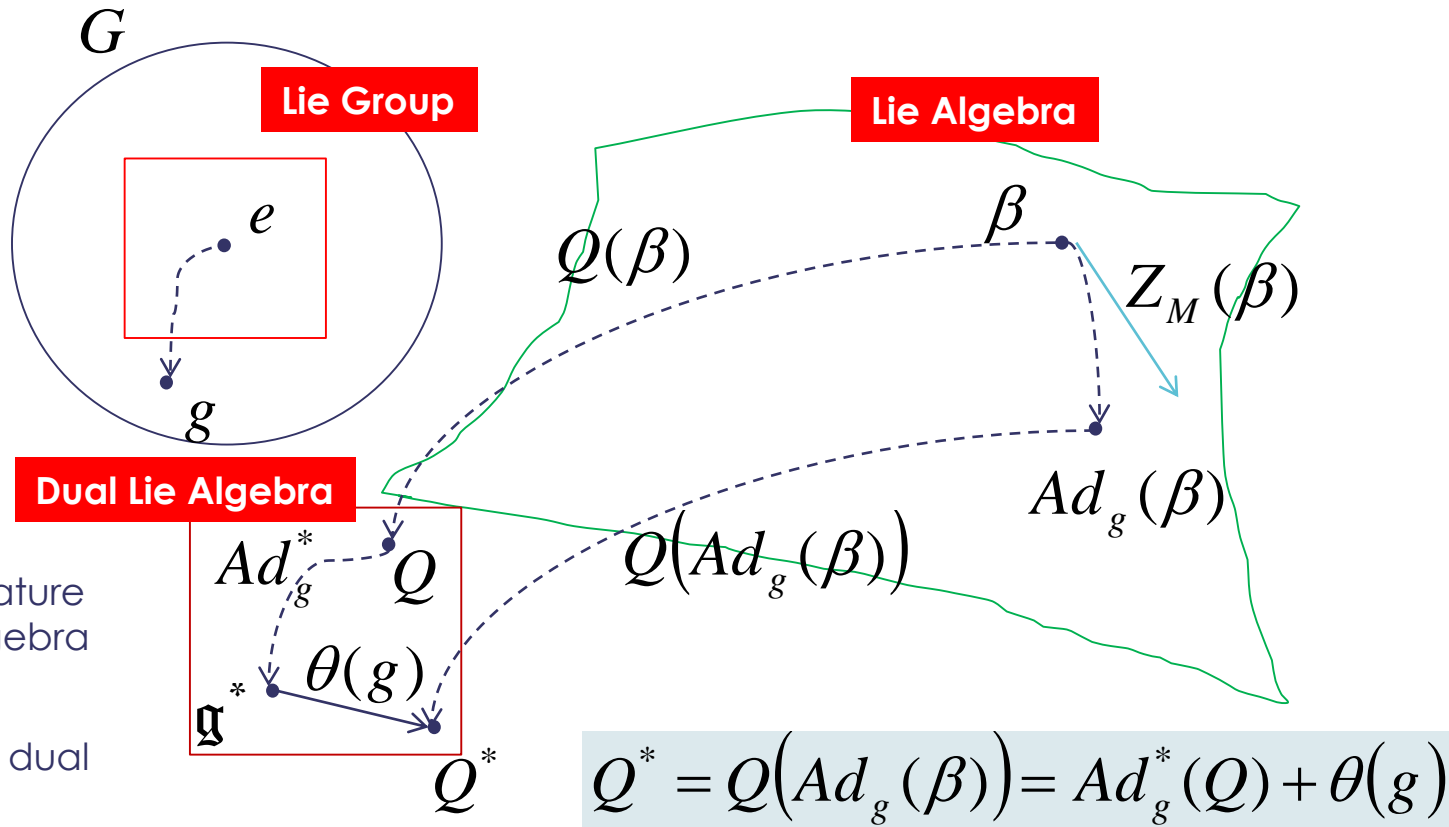
- In Souriau's **Lie groups thermodynamics**, the invariance by re-parameterization in information geometry has been replaced by invariance with respect to the action of the group. When an element of the group g acts on the element $\beta \in \mathfrak{g}$ of the Lie algebra, given by adjoint operator Ad_g . Under the action of the group, $Ad_g(\beta)$, **the entropy $S(Q)$ and the Fisher metric $I(\beta)$ are invariant:**

$$\beta \in \mathfrak{g} \rightarrow Ad_g(\beta) \Rightarrow \begin{cases} S[Q(Ad_g(\beta))] = S(Q) \\ I[Ad_g(\beta)] = I(\beta) \end{cases}$$

$$I(\beta) = -\frac{\partial^2 \Phi}{\partial \beta^2} \text{ with } \Phi(\beta) = -\log \int_M e^{-\langle \beta, U^{(\xi)} \rangle} d\lambda$$

$$S(Q) = \langle \beta, Q \rangle - \Phi(\beta) \text{ with } Q = \frac{\partial \Phi(\beta)}{\partial \beta} \in \mathfrak{g}^* \text{ and } \beta = \frac{\partial S(Q)}{\partial Q} \in \mathfrak{g}$$

Fundamental Souriau Theorem



Fisher-Souriau Metric as a non-null Cohomology extension of KKS 2 form (Kirillov-Kostant-Souriau 2 form)

■ Souriau definition of Fisher Metric is related to the extension of KKS 2-form (Kostant-Kirillov-Souriau) in case of non-null Cohomogy:

Souriau-Fisher Metric

$$I(\beta) = [g_\beta] \quad \text{with} \quad g_\beta([\beta, Z_1], [\beta, Z_2]) = \tilde{\Theta}_\beta(Z_1, [\beta, Z_2])$$

$$\text{with } \tilde{\Theta}_\beta(Z_1, Z_2) = \tilde{\Theta}(Z_1, Z_2) + \langle Q, [Z_1, Z_2] \rangle$$

Non-null cohomology: additional term from Souriau Cocycle

Equivariant KKS 2 form

$$\tilde{\Theta}(X, Y) = J_{[X, Y]} - \{J_X, J_Y\} \quad \text{with} \quad J(x): M \rightarrow \mathfrak{g}^* \quad \text{such that} \quad J_X(x) = \langle J(x), X \rangle, \quad X \in \mathfrak{g}$$

$$\tilde{\Theta}(X, Y): \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{R}$$

$$\text{with } \Theta(X) = T_e \theta(X(e))$$

$$\tilde{\Theta}(\beta, Z) + \langle Q, [\beta, Z] \rangle = 0$$

$$X, Y \mapsto \langle \Theta(X), Y \rangle$$

$$\beta \in \text{Ker } \tilde{\Theta}_\beta$$

**Souriau Fundamental
Equation of Lie Group Thermodynamics**

$$Q(Ad_g(\beta)) = Ad_g^*(Q) + \theta(g)$$

Souriau-Fisher Metric & Souriau Lie Groups Thermodynamics: Bedrock for Lie Group Machine Learning

TEMPERATURE
In Lie Algebra

HEAT
In Dual Lie Algebra

Fisher Metric

$$g_{\beta}([\beta, Z_1], [\beta, Z_2]) = \tilde{\Theta}_{\beta}(Z_1, [\beta, Z_2]) \geq 0$$

$$I(\beta) = I(Ad_g(\beta)) = -\frac{\partial^2 \Phi}{\partial \beta^2} = \frac{\partial^2 \log \int e^{-\langle \beta, U(\xi) \rangle} d\lambda}{\partial \beta^2}$$

Entropy invariant under the action of the group

Legendre

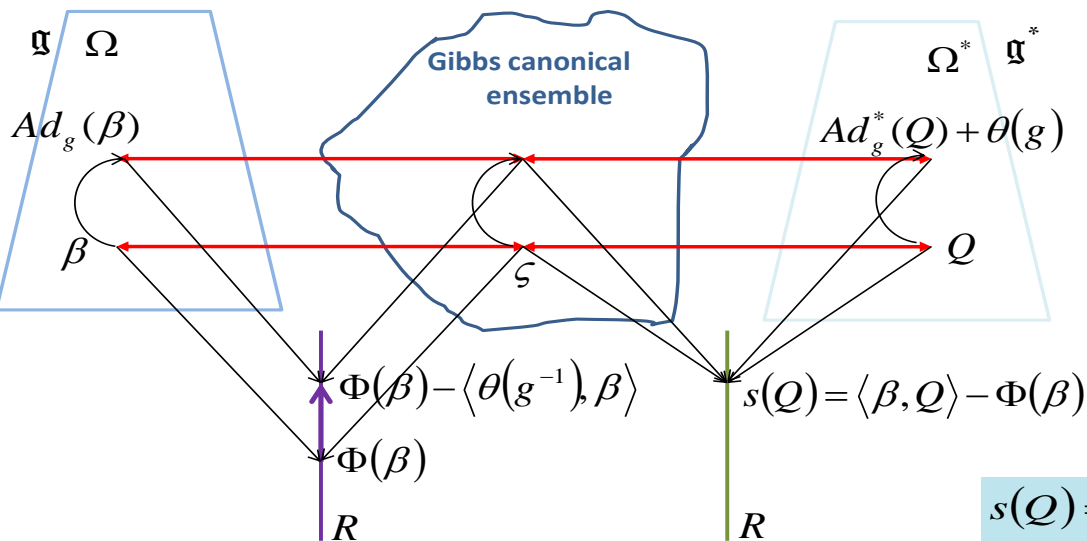
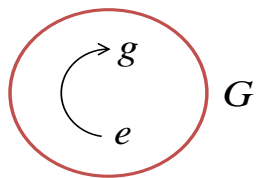
Clairaut

$$s(Q) = \langle \beta, Q \rangle - \Phi(\beta) = \langle \Theta^{-1}(Q), Q \rangle - \Phi(\Theta^{-1}(Q))$$

$$Q = \Theta(\beta) = \frac{\partial \Phi}{\partial \beta} \in \mathfrak{g}^*$$

$$\beta = \Theta^{-1}(Q) \in \mathfrak{g}$$

THALES



Logarithm of Partition Function
(Massieu Characteristic Function)

Entropy

Covariant Souriau-Gibbs density

- Souriau has then defined a Gibbs density that is covariant under the action of the group:

$$p_{Gibbs}(\xi) = e^{\Phi(\beta) - \langle U(\xi), \beta \rangle} = \frac{e^{-\langle U(\xi), \beta \rangle}}{\int_M e^{-\langle U(\xi), \beta \rangle} d\lambda_\omega}$$

$$\text{with } \Phi(\beta) = -\log \int_M e^{-\langle U(\xi), \beta \rangle} d\lambda_\omega$$

$$Q = \frac{\partial \Phi(\beta)}{\partial \beta} = \frac{\int_M U(\xi) e^{-\langle U(\xi), \beta \rangle} d\lambda_\omega}{\int_M e^{-\langle U(\xi), \beta \rangle} d\lambda_\omega} = \int_M U(\xi) p(\xi) d\lambda_\omega$$

- We can express the Gibbs density with respect to Q by inverting the relation

$$Q = \frac{\partial \Phi(\beta)}{\partial \beta} = \Theta(\beta) . \text{ Then } p_{Gibbs, Q}(\xi) = e^{\Phi(\beta) - \langle U(\xi), \Theta^{-1}(Q) \rangle} \text{ with } \beta = \Theta^{-1}(Q)$$

Souriau Entropy Invariance

Casimir Invariant Function in coadjoint representation

- We observe that Souriau Entropy $S(Q)$ defined on coadjoint orbit of the group has a property of invariance :

$$S(Ad_g^\#(Q)) = S(Q)$$

- with respect to Souriau affine definition of coadjoint action:

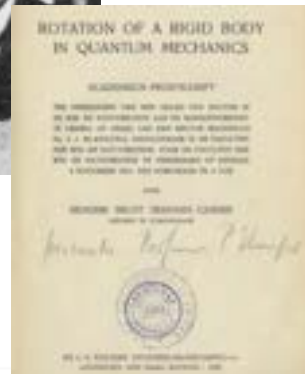
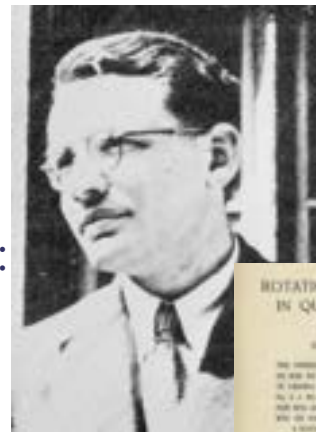
$$Ad_g^\#(Q) = Ad_g^*(Q) + \theta(g)$$

- where $\theta(g)$ is called the Souriau cocycle.

$$Q(Ad_g(\beta)) = Ad_g^*(Q) + \theta(g)$$

$$S(Q(Ad_g(\beta))) = S(Q)$$

New Entropy Definition:
Casimir Function in
Coadjoint Representation
Invariant under the action
of the Group



Hendrik Casimir
(Thesis supervised by
Niels Bohr & Paul Ehrenfest)

**H.B.G. Casimir, On the Rotation of a Rigid Body in
Quantum Mechanics, Doctoral Thesis, Leiden, 1931.**

Entropy as Invariant Casimir Function in Coadjoint Representation

NEW GEOMETRIC DEFINITION OF ENTROPY

$$\{S, H\}_{\tilde{\Theta}}(Q) = 0$$

$$ad_{\frac{\partial S}{\partial Q}}^* Q + \Theta\left(\frac{\partial S}{\partial Q}\right) = 0$$

$$\{S, H\}(Q) = \left\langle Q, \left[\frac{\partial S}{\partial Q}, \frac{\partial H}{\partial Q} \right] \right\rangle = -C_{ij}^k Q_k \frac{\partial S}{\partial Q_i} \cdot \frac{\partial H}{\partial Q_j}$$

$$[e_i, e_j] = C_{ij}^k e_k, \quad C_{ij}^k \text{ structure coefficients}$$

$$\{S, H\}_{\tilde{\Theta}}(Q) = \left\langle Q, \left[\frac{\partial S}{\partial Q}, \frac{\partial H}{\partial Q} \right] \right\rangle + \tilde{\Theta}\left(\frac{\partial S}{\partial Q}, \frac{\partial H}{\partial Q}\right) = 0, \quad \forall H: \mathfrak{g}^* \rightarrow \mathbb{R}, \quad Q \in \mathfrak{g}^*$$

$$\tilde{\Theta}(X, Y) = J_{[X, Y]} - \{J_X, J_Y\} \quad \text{where } J_X(x) = \langle J(x), X \rangle$$

$$\tilde{\Theta}(X, Y) = \langle \Theta(X), Y \rangle \quad \text{with } \Theta(X) = T_e \theta(X(e))$$

$$\theta(g) = Q(Ad_g(\beta)) - Ad_g^*(Q)$$

Geometric Poincaré-Souriau-Fourier Heat Equation

Poincaré-Souriau-Fourier heat equation

$$\frac{\partial Q}{\partial t} = ad_{\frac{\partial H}{\partial Q}}^* Q + \Theta \left(\frac{\partial H}{\partial Q} \right) = \{Q, H\}_{\tilde{\Theta}} \xrightarrow{Q = \frac{\partial \Phi}{\partial \beta}} \frac{d}{dt} \frac{\partial \Phi}{\partial \beta} = ad_{\frac{\partial H}{\partial Q}}^* \frac{\partial \Phi}{\partial \beta} + \Theta \left(\frac{\partial H}{\partial Q} \right)$$

Link with Classical Fourier Heat Equation

- Heat Equation is the PDE for (calorific) Energy density where the nature of material is characterized by the geometric heat capacity. In the Euclidean case with homogeneous material, we recover classical equation on density of Energy ρ_E

$$\frac{\partial \rho_E}{\partial t} = \operatorname{div} \left(\frac{\lambda}{C} \nabla \rho_E \right) \text{ with } \frac{\partial \rho_E}{\partial t} = C \frac{\partial T}{\partial t}$$

Balian, R., Introduction à la thermodynamique hors-équilibre. 2003, CEA report

Motivation for Lie Group Machine Learning: Data in Homogenous Space where a Lie Groups act homogeneously

(Micro-)Doppler & Space-Time wave Learning in Poincaré/Siegel Polydisks

$$SU(1,1) = \left\{ \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} / \alpha, \beta \in \mathbb{C}, |\alpha|^2 - |\beta|^2 = 1 \right\}$$

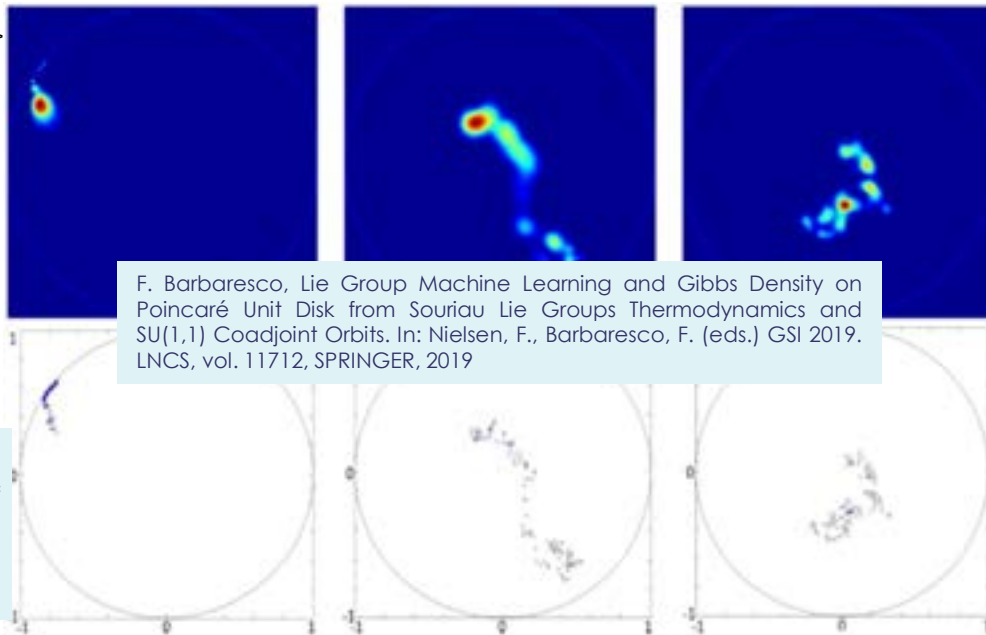
$$\varphi: THDP(n) \rightarrow R_+^* \times D^{n-1}$$

$$R_n \mapsto (P_0, \mu_1, \dots, \mu_{n-1})$$

$$R_n = (h, \mu_1, \dots, \mu_{n-1}) \in R^{+*} \times D^{n-1}$$

$$\text{with } (D^{n-1} = D \times \dots \times D)$$

$$ds_{\text{Poincaré}}^2 = \frac{|dz|^2}{(1-|z|^2)^2} = (1-zz^*)^{-1} dz (1-z^*z)^{-1} dz^*$$



Matrix Lie Group $SU(1,1)$ for Doppler Data

Poincaré Unit Disk as a Homogeneous Manifold

- The Poincaré unit disk is an homogeneous bounded domain where the Lie Group $SU(1,1)$ act transitively. This Matrix Group is given by

$$SU(1,1) = \left\{ \begin{bmatrix} a & b \\ b^* & a^* \end{bmatrix} / |a|^2 - |b|^2 = 1, a, b \in \mathbb{C} \right\}$$

- where $SU(1,1)$ acts on the Poincaré Unit Disk by: $g \in SU(1,1) \Rightarrow g.z = \frac{az + b}{b^*z + a^*}$
- with Cartan Decomposition of $SU(1,1)$:

$$\begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} = |a| \begin{pmatrix} 1 & z \\ z^* & 1 \end{pmatrix} \begin{pmatrix} a/|a| & 0 \\ 0 & a^*/|a| \end{pmatrix}$$

$$\text{with } z = b(a^*)^{-1}, |a| = (1 - |z|^2)^{-1/2}$$

- We can observe that $z = b(a^*)^{-1}$ could be considered as action of $g \in SU(1,1)$ on the centre on the unit disk $z = g.0 = b(a^*)^{-1}$.

Poincaré Unit Disk and $SU(1,1)$ Lie Group

► The group of complex unimodular pseudo-unitary matrices $SU(1,1)$:

$$G = SU(1,1) = \left\{ \begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} / |a|^2 - |b|^2 = 1, a, b \in \mathbb{C} \right\}$$

► the Lie algebra $\mathfrak{g} = \mathfrak{su}(1,1)$ is given by:

$$\mathfrak{g} = \left\{ \begin{pmatrix} -ir & \eta \\ \eta^* & ir \end{pmatrix} / r \in \mathbb{R}, \eta \in \mathbb{C} \right\}$$

with the following bases $(u_1, u_2, u_3) \in \mathfrak{g}$:

$$u_1 = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, u_2 = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, u_3 = \frac{1}{2} \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$$

with the commutation relation:

$$[u_3, u_2] = u_1, [u_3, u_1] = u_2, [u_2, u_1] = -u_3$$

Poincaré Unit Disk and $SU(1,1)$ Lie Group

➤ Dual base on dual Lie algebra is named

$$(u_1^*, u_2^*, u_3^*) \in \mathfrak{g}^*$$

➤ The dual vector space $\mathfrak{g}^* = \mathfrak{su}^*(1,1)$ can be identified with the subspace of $\mathfrak{sl}(2, \mathbb{C})$ of the form:

$$\mathfrak{g}^* = \left\{ \begin{pmatrix} z & x+iy \\ -x+iy & -z \end{pmatrix} = x \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} + z \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} / x, y, z \in \mathbb{R} \right\}$$

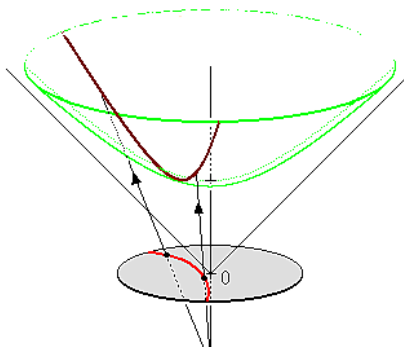
➤ Coadjoint action of $g \in G$ on dual Lie algebra $\xi \in \mathfrak{g}^*$ is written $g.\xi$

Coadjoint Orbit of $SU(1,1)$ and Souriau Moment Map

$$J(z) = r \left(\frac{z + z^*}{(1 - |z|^2)} u_1^* + \frac{z - z^*}{i(1 - |z|^2)} u_2^* + \frac{1 + |z|^2}{(1 - |z|^2)} u_3^* \right) \in O(ru_3^*), z \in D$$

- J is linked to the natural action of G on D (by fractional linear transforms) but also the coadjoint action of G on $O(ru_3^*) = G/K$
- J^{-1} could be interpreted as the stereographic projection from the two-sphere S^2 onto $C \cup \infty$:

The coadjoint action of $G = SU(1,1)$ is the upper sheet $x_3 > 0$ of the two-sheet hyperboloid



Charles-Michel Marle, Projection stéréographique et moments, hal-02157930, version 1, Juin 2019

184 $\left\{ \xi = x_1 u_1^* + x_2 u_2^* + x_3 u_3^* : -x_1^2 - x_2^2 + x_3^2 = r^2 \right\}$

Invariant Moment Map

- The associated moment map $J : D \rightarrow su^*(1,1)$ defined by $J(z).u_i = J_i(z, z^*)$, maps D into a coadjoint orbit in $su^*(1,1)$.
- Then, we can write the moment map as a matrix element of $su^*(1,1)$:

$$J(z) = J_1(z, z^*)u_1^* + J_2(z, z^*)u_2^* + J_3(z, z^*)u_3^*$$

$$J(z) = \rho \begin{pmatrix} \frac{1+|z|^2}{1-|z|^2} & -2\frac{z^*}{1-|z|^2} \\ 2\frac{z}{1-|z|^2} & -\frac{1+|z|^2}{1-|z|^2} \end{pmatrix} \in \mathfrak{g}^*$$

$$\mathfrak{g}^* = \left\{ \begin{pmatrix} z & x+iy \\ -x+iy & -z \end{pmatrix} = x \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} + z \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} / x, y, z \in \mathbb{R} \right\}$$

$$p_{Gibbs}(z) = \frac{e^{-\left\langle \rho \begin{pmatrix} \frac{1+|z|^2}{(1-|z|^2)} & \frac{-2z^*}{(1-|z|^2)} \\ \frac{2z}{(1-|z|^2)} & -\frac{1+|z|^2}{(1-|z|^2)} \end{pmatrix}, \begin{pmatrix} ir & \eta \\ \eta^* & -ir \end{pmatrix} \right\rangle}}{\int_D e^{-\langle J(z), \beta \rangle} d\lambda(z)}$$

- To write the Gibbs density with respect to its statistical moments, we have to express the density with respect to $Q = E[J(z)]$
- Then, we have to invert the relation between Q and β , to replace by $\beta = \Theta^{-1}(Q) \in \mathfrak{g}$ where $Q = \frac{\partial \Phi(\beta)}{\partial \beta} = \Theta(\beta) \in \mathfrak{g}^*$ with $\Phi(\beta) = -\log \int_D e^{-\langle J(z), \beta \rangle} d\lambda(z)$ deduce from Legendre transform. The mean moment map is given by:

$$Q = E[J(z)] = E \left[\rho \begin{pmatrix} \frac{1+|w|^2}{(1-|w|^2)} & \frac{-2w^*}{(1-|w|^2)} \\ \frac{2w}{(1-|w|^2)} & -\frac{1+|w|^2}{(1-|w|^2)} \end{pmatrix} \right] \quad \text{where } w \in D$$

Gauss Density on Siegel Unit Disk

Moment Map of $SU(n,n)/S(U(n) \times U(n))$

➤ The moment map for $SU(n,n)/S(U(n) \times U(n))$ is then given by:

$$J(Z) = \rho n \begin{pmatrix} (I_n - ZZ^+)^{-1} (I_n + ZZ^+) & -2Z^+ (I_n - ZZ^+)^{-1} \\ 2(I_n - ZZ^+)^{-1} Z & (I_n + ZZ^+) (I_n - ZZ^+)^{-1} \end{pmatrix} \in \mathfrak{g}^*$$

➤ The Souriau Gibbs density is then given with $\beta, M \in \mathfrak{g}$ and $Z \in SD_n$ by:

$$p_{\text{Gibbs}}(Z) = \frac{e^{-\left\langle \rho n \begin{pmatrix} (I_n - ZZ^+)^{-1} (I_n + ZZ^+) & -2Z^+ (I_n - ZZ^+)^{-1} \\ 2(I_n - ZZ^+)^{-1} Z & (I_n + ZZ^+) (I_n - ZZ^+)^{-1} \end{pmatrix}, \beta \right\rangle}}{\int_{SD_n} e^{-\langle J(Z), \beta \rangle} d\lambda(Z)}$$

$$\beta = \Theta^{-1}(Q) \in \mathfrak{g}$$

$$Q = E[J(Z)]$$

$$Q = \frac{\partial \Phi(\beta)}{\partial \beta} = \Theta(\beta) \in \mathfrak{g}^*$$

➤ Gauss density of SPD matrix is given by Cayley Transform with:

$$Z = (Y - I)(Y + I)^{-1}_{\text{OPEN}}, Y \in \text{Sym}(n)^+$$

Pathological Clutter Segmentation by Supervised Classification: Lie-Group Equivariant Neural Networks: G-CNN for $SU(1,1)$ Lie Group

Ongoing internal work for G-CNN application to radar signal processing

- Represent the Doppler signals as complex covariance matrices
- Embed these matrices into the Poincaré poly-disk (hyperbolic space)
- Build equivariant convolution kernels to $SU(1,1)$
- Hope to improve robustness to sensor noise

$$SU(1,1) = \left\{ g_{\alpha,\beta} = \begin{bmatrix} \alpha & \beta \\ \beta & \bar{\alpha} \end{bmatrix}, |\alpha|^2 - |\beta|^2 = 1, \alpha, \beta \in \mathbb{C} \right\}$$

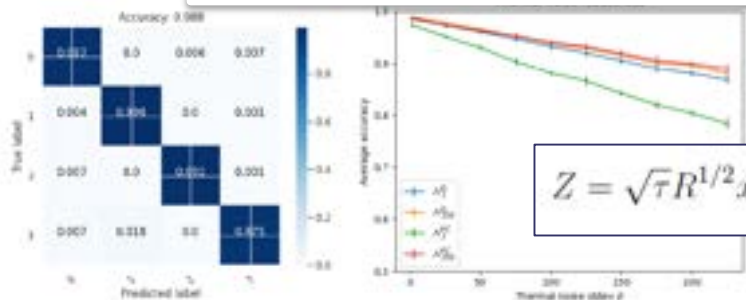
$SU(1,1)$ acts on the Poincaré Disk by Mobius transform

$$g_{\alpha,\beta} \circ z = \frac{\alpha z + \beta}{\beta z + \bar{\alpha}}$$

$$U(1) = \left\{ \begin{bmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{bmatrix}, |\alpha|^2 = 1, \alpha, \beta \in \mathbb{C} \right\}$$

P.Y. Lagrave, Y. Cabanes, F. Barbaresco, An Equivariant Neural Network with Hyperbolic Embedding for Robust Doppler Signal Classification, International Radar Symposium, IRS'21, Berlin, June 2021

Pierre-Yves Lagrave, TRT



$$Z = \sqrt{\tau} R^{1/2} x + b_{\text{radar}}$$

| T_σ | N_1^G | N_1^{FC} | N_{250}^G | N_{250}^{FC} |
|------------|-----------------|-----------------|-----------------|-----------------|
| T_1 | 0.9873 (0.0015) | 0.9735 (0.0017) | 0.9891 (0.0021) | 0.9846 (0.0026) |
| T_{50} | 0.9606 (0.0039) | 0.9301 (0.0054) | 0.9653 (0.0034) | 0.9621 (0.0047) |
| T_{100} | 0.9327 (0.0074) | 0.8811 (0.0041) | 0.9386 (0.0054) | 0.9401 (0.0079) |
| T_{150} | 0.9051 (0.0054) | 0.8419 (0.0044) | 0.9148 (0.0032) | 0.9191 (0.0056) |
| T_{200} | 0.8805 (0.0066) | 0.8042 (0.0038) | 0.8955 (0.0066) | 0.8988 (0.0043) |

OPEN

THALES

QUESTIONS: C.R. Rao and SPRINGER Information Geometry Journal

EDITORIAL



Congratulatory message

Calyampudi Radhakrishna Rao¹

Published online: 19 September 2018
© Springer Nature Singapore Pte Ltd. 2018

I am glad to know that Springer is starting a new journal with the title of *Information Geometry* under the chief editorship of Shinto Eguchi with co-editors Nihat Ay, Frank Nielsen, and Jun Zhang. The journal is interdisciplinary, integrating various disciplines, especially branches of mathematical sciences related to the field of information geometry. This is a needed area of literature, and the journal meets that requirement.

Congratulations and best wishes for the success of the journal.



C. R. Rao

C.R. Rao

Information Geometry

Volume 4 - Number 1 - 2021

Editor in Chief
Nihat Ay (Hamburg)

Co-Editors
Shinto Eguchi (Tokyo)
Hirosi Matsuzoe (Nagoya)
Frank Nielsen (Tokyo)
Jun Zhang (Jilin)

Associate Editors
Toshiaki Banerjee (Paris)
Damiano Braga (London)
Dagmar C. Braddy (London)
Shiro Noda (Tokyo)
Hiroshi Nishitani (Tokyo)
Paul Norvaiša (Ljubljana)
Frankfurt Maru (Prague)
Mitsuru Morita (Tokyo)

Hirosi Nagasaki (Tokyo)
Jan Naudts (Antwerp)
Nigel Newton (Canberra)
Richard Nickl (Canberra)
Atsushi Ohara (Tokyo)
Giovanni Pinzone (Turin)
Constantine Tsallis (Rio de Janeiro)

Advisory Board
Ole E. Barndorff-Nielsen (Aarhus)
David Cox (Oxford)
Bradley Efron (Stanford)
C.R. Rao (Hyderabad)

Honorary Editors
Shun-ichi Amari (Tokyo)
Imre Csiszar (Budapest)

Springer