Anastasis Kratsios

McMaster University
Department of Mathematics and Statistics
Hamilton Hall, Hamilton
Canada

Foundations of Geometric Deep Learning

Specialization: I investigate the analytic and statistical foundations of deep learning. By understanding the behaviour of these models, I can rigorously modify them to efficiently encode the structures arising in the natural and social sciences, esp. stochastic processes, games, and PDEs.

Primary Areas: Mathematical and statistical foundations of (geometric deep and operator) learning. **Applied Areas:** Deep learning for optimal control, game theory, PDEs, and finance.

Fmp	loyment	History
	I O y I I I C I I C	1 113601 9

Academic

2022-Current Assistant Professor - Data Science in Mathematical Finance, McMaster University, Hamilton, Canada
Cross Appointment: Computational Science and Engineering

2023-Current Faculty Affiliate, Vector Institute, Toronto, Canada

2021-2022 **Postdoctoral Researcher - Geometric Deep Learning**, *University of Basel*, Basel, Switzerland, Groups: Computer Science and Probability

2018-2021 **Postdoctoral Researcher - Stochastic Finance and Geometric Deep Learning**, *ETH Zürich*, Zürich, Switzerland, Group: Mathematical Finance
Private Sector

2017 Risk Management Researcher, CIBC Capital Markets, Toronto, Canada

Education

- 2014–2018 **PhD Mathematics (Mathematical Finance and Analysis)**, Concordia University, Montreal, Supersivors: Cody Hyndman and Alina Stancu
- 2012–2014 **MSc Mathematics (Homological Algebra and Non-Commutative Geometry)**, *University of Montreal*, Montreal, Supervisor: Abraham Broer
- 2009–2012 BA Mathematics (Pure Mathematics), Concordia University, Montreal

Grants

- 2023-2028 **NSERC Discovery Grant**, \$205,000 Foundations of Geometric Deep Learning and Trustworthy AI in Finance
- 2023-2028 NSERC Early Career Grant, \$12,500
 - 2023 McMaster Startup Grant, \$130,000

Research Awards and Recognitions

- 2024 **NeurIPS Spotlight Paper**, Energy-Guided Continuous Entropic Barycenter Estimation for General Costs
- 2022 **ICLR Spotlight Paper**, *Universal Approximation Under Constraints is Possible with Transformers*

Research Articles (40)

Highlights:

- 4 In The Journal of Machine Learning Research (JMLR)
- 9 In NeurIPS, ICLR, ICML, COLT, or TMLR
- 17 Productivity in 2024: New Preprints and Accepted Manuscripts

Legend:

Each research article will be classified into one of the following topics.

- Finance, 5
- Control, Games, and Optimization, 3
- Time-Series and Stochastic Processes, 5
- Statistics and Learning Theory, 7
- Deep Learning Theory, (Approximation, Interpolation, and Embedding), 18
- Deep Learning in PDEs and Inverse Problems, 2
- Other, 2

Publications (24)

Each student or postdoc I mentored is emphasized with an $(^{\dagger})$ symbol.

- 2024 Cheng[†], T. S., Lucchi, A., Kratsios, A., and Belius, D., *A Comprehensive Analysis on the Learning Curve in Kernel Ridge Regression*, NeurIPS (2024)
 H5-index: 337

 Statistics and Learning Theory
- Kolesov, A., Mokrov, P., Udovichenko, I., Gazdieva, M., Pammer, G., A.
 Kratsios, Korotin, A., and Burnaev, E., Energy-Guided Continuous Entropic
 Barycenter Estimation for General Costs, NeurIPS (2024), Award: Spotlight
 H5-index: 337
 Deep Learning Theory
 Statistics and Learning Theory
- 2024 **Cheng**[†], **T. S., Lucchi, A., Kratsios, A., and Belius, D.**, *Characterizing overfitting in kernel ridgeless regression through the eigenspectrum*, International Conference on Learning (2024)

H5-index: 304

• Statistics and Learning Theory

- 2024 Benitez[†], J. A. L., Furuya, T., Faucher, F., Kratsios, A., Tricoche, X., and de Hoop, M. V., Out-of-distributional risk bounds for neural operators with applications to the Helmholtz equation. Journal of Computational Physics, 113168.

 H5-index: 89

 PDEs and Inverse Problems Statistics and Learning Theory
- 2024 Kratsios, A., Hong[†], R., and de Ocáriz Borde, H., Capacity bounds for hyperbolic neural network representations of latent tree structures. Neural Networks, 106420.

 H5-index: 95

 Deep Learning Theory
- 2023 Kassraie, P., Hou, S., Kratsios, A., Krause, A., and Rothfuss, J., Instance-Dependent Generalization Bounds via Optimal Transport. Journal of Machine Learning Research, 24, 349.

H5-index: 117

• Statistics and Learning Theory

- 2023 Cheng, T. S., Kratsios, A., Lucchi, A., Dokmanic, I., and Belius, D., A theoretical analysis of the test error of finite-rank kernel ridge regression. Advances in Neural Information Processing Systems, 36, 4767-4798.
 - H5-index: 337

 Statistics and Learning Theory
- 2023 Kratsios, A., Debarnot, V., and Dokmanić I., Small transformers compute universal metric embeddings. Journal of Machine Learning Research, 24(170), 1-48.
 H5-index: 117
 Deep Learning Theory
- 2023 Acciaio, B., Kratsios, A., and Pammer, G., Designing universal causal deep learning models: The geometric (hyper) transformer, Mathematical Finance Special Issue on Machine Learning in Finance

 H5-index: 82

 Finance
- 2023 Kratsios, A., Universal regular conditional distributions via probabilistic transformers.
 Constructive Approximation, 57(3), 1145-1212.
 H5-index: 54
 Deep Learning Theory
- 2023 Herrera, C., Krach, F., Kratsios, A., Ruyssen, P., and Teichmann, J., Denise:
 Deep Robust Principal Component Analysis for Positive Semidefinite Matrices.
 Transactions on Machine Learning Research.

 H5-index: NA

 Deep Learning Theory
- 2022 **Papon**[†], **L., and Kratsios, A.**, *Universal approximation theorems for differentiable geometric deep learning. Journal of Machine Learning Research, 23(196), 1-73.*H5-index: 117

 Deep Learning Theory
- 2023 **Hyndman, C., and Kratsios, A.**, Generative Ornstein-Uhlenbeck Markets via Geometric Deep Learning. In International Conference on Geometric Science of Information (pp. 605-614). Springer Nature Switzerland.

 H5-index: NA

 Finance
- 2022 Kratsios, A., Zamanlooy[†], B., Dokmanic, I., and Liu[†], T., Universal approximation under constraints is possible with transformers. In ICLR) International Conference on Learning Representations., Award: Spotlight
 H5-index: 304

 Deep Learning Theory
- Zamanlooy[†], B., and Kratsios, A., Do ReLU Networks Have An Edge When Approximating Compactly-Supported Functions?. Transactions on Machine Learning Research.
 H5-index: NA
 ▶ Deep Learning Theory
- 2022 **Zamanlooy**[†], **B., and Kratsios, A.**, Learning sub-patterns in piecewise continuous functions. Neurocomputing, 480, 192-211.

 H5-index: 136

 Deep Learning Theory
- 2021 **Hyndman, C., and Kratsios, A.**, *NEU: A meta-algorithm for universal UAP-invariant feature representation. Journal of Machine Learning Research, 22(92), 1-51.*
- H5-index: 117 Deep Learning Theory
- Casgrain, P., and Kratsios, A., Optimizing Optimizers: Regret-optimal gradient descent algorithms. In Conference on Learning Theory (pp. 883-926). PMLR.
 H5-index: 54
 Control, Games, and Optimization

2021 **Kratsios, A.**, Lower-estimates on the Hochschild (Co)homological dimension of commutative algebras and applications to smooth affine schemes and quasi-free algebras. Mathematics, 9(3), 251.

H-index: 43

2021 **Kratsios, A.**, The universal approximation property: characterization, construction, representation, and existence. Annals of Mathematics and Artificial Intelligence, 89(5), 435-469.

H-index: 59 • Deep Learning Theory

- 2020 Kratsios, A., Bilokopytov, I., Non-Euclidean universal approximation. Advances in Neural Information Processing Systems, 33, 10635-10646.
 H5-index: 337
 Deep Learning Theory
- 2020 **Kratsios, A., Hyndman, C.**, Deep arbitrage-free learning in a generalized HJM framework via arbitrage-regularization. Risks, 8(2), 40.

 H-index: 22

 Finance
- 2020 **Kratsios, A., Hyndman, C., Wang, R.**, *The entropic measure transform. Canadian Journal of Statistics, 48(1), 97-129.*H-index: 53 Finance

Under Review (15)

- 2024 **G. A. Alvarez, I. Ekren, A. Kratsios, X. Yang**, Neural Operators Can Play Dynamic Stackelberg Games, ArXiV: 2411.09644
 - Control, Games, and Optimization
 Deep Learning Theory
- 2024 Borde, H., Lukoianov, A., Kratsios, A., Law, M. T., Dong, X., and Bronstein, M., Scalable Message Passing Neural Networks: No Need for Attention in Large Graph Representation Learning, ArXiV: 2408.13885
 - Deep Learning Theory
- 2024 Borde, H., Kratsios, A., Law, M. T., Dong, X., and Bronstein, M., Neural Spacetimes for DAG Representation Learning, ArXiV: 2408.13885
 - Deep Learning Theory
- 2024 Persiianov, M., Asadulaev, A., Andreev, N., Starodubcev, N., Baranchuk, D., Kratsios, A., Burnaev, E. and Korotin, A., Inverse Entropic Optimal Transport Solves Semi-supervised Learning via Data Likelihood Maximization, ArXiV:2410.02628
 - PDEs and Inverse Problems
- 2024 Hong[†], R. and Kratsios, A., Bridging the Gap Between Approximation and Learning via Optimal Approximation by ReLU MLPs of Maximal Regularity, ArXiV: 2409.12335
 - Deep Learning Theory
- 2024 Kratsios, A., Sáez de Ocáriz Borde, H., Furuya, T., and Law, M. T., Approximation Rates and VC-Dimension Bounds for (P)ReLU MLP Mixture of Experts. ArXiV: ArXiV-2402.
 - Deep Learning Theory

2024 Arabpour[†], R., Armstrong, J., Galimberti, L., Kratsios, A., and Livieri, G., Low-dimensional approximations of the conditional law of Volterra processes: a non-positive curvature approach. ArXiV preprint ArXiV:2405.20094., Revision - Constructive Approximation

• Time-Series and Causality

2024 Limmer[†], Y., Kratsios, A., Yang[†], X., Saqur, R., and Horvath, B., Reality Only Happens Once: Single-Path Generalization Bounds for Transformers. ArXiV preprint ArXiV:2405.16563.

• Time-Series and Causality

2024 Kratsios, A., Furuya, T., Lara Benitez[†], J.A., Lassas, M., and de Hoop, M., Mixture of Experts Soften the Curse of Dimensionality in Operator Learning. ArXiV preprint ArXiV:2404.09101.

• Deep Learning Theory

2024 Kratsios, A., Neuman, A. M., and Pammer, G., Tighter Generalization Bounds on Digital Computers via Discrete Optimal Transport. ArXiV preprint ArXiV:2402.05576, (Revision Transactions of Information Theory)

• Statistics and Learning Theory

2024 Saqur, R., Kratsios, A., Krach[†], F., Limmer[†], Y., Tian, J. J., Willes, J., Horvarth, B. and Rudzicz, F. (2024)., Filtered not Mixed: Stochastic Filtering-Based Online Gating for Mixture of Large Language Models. ArXiV preprint ArXiV:2406.02969.

• Control, Games, and Optimization • Finance

2023 **Hovart, B., Kratsios, A., Limmer[†], Y., and Yang[†], X.**, Deep Kalman Filters Can Filter. ArXiV preprint ArXiV:2310.19603.

• Time-Series and Causality

2023 Kratsios, A., Liu, C., Lassas, M., de Hoop, M. V., and Dokmanić, I., An Approximation Theory for Metric Space-Valued Functions With A View Towards Deep Learning. ArXiV preprint ArXiV:2304.12231.

• Time-Series and Causality

2023 Yang[†], X., Kratsios, A., Krach, F., Grasselli, M., and Lucchi, A., Regret-Optimal Federated Transfer Learning for Kernel Regression with Applications in American Option Pricing. ArXiV preprint ArXiV:2309.04557.

• Control, Games, and Optimization • Finance

2022 **Galimberti, L., Kratsios, A., and Livieri, G.**, Designing universal causal deep learning models: The case of infinite-dimensional dynamical systems from stochastic analysis. ArXiV preprint ArXiV:2210.13300.

• Time-Series and Causality

Other Publications (1)

2024 **B. Horvath, A. Kratsios, R. Saqur**, Book Review: Book Deep Learning - Foundations and Concepts of Christopher M. Bishop and Hugh Bishop, Springer Nature

Other

Supervision

Postdoctoral Researcher(s)

- 2022-2025 **Postdoctoral Researcher Optimal Control**, *Xuwei Yang*, McMaster University, Learning how to control and filter (Joint: Matheus Grasselli)

 PhD Student(s)
- 2024-2028 M.Sc. to Ph.D. Math. Candidate, Hossein Rouhvarzi, McMaster University, Intrinsic approximation in stochastic processes and applications to finance.

 Officially Transfers to PhD in Summer 2025
- 2023-2027 **Ph.D. Stats. Candidate Joint: Paul McNicholas**, *Yicen Li*, McMaster University, Foundation models for deep clustering.
- 2023-2027 **Ph.D. Math. Candidate**, *Ruiyang Hong*, McMaster University, Neural Extension Theorems Approximation by data-interpolating MLPs of maximal regularity .

 Masters Student(s)
- 2023-2025 **M.Sc. CSE**, *Reza Arabpour*, McMaster University, Implementations of hypernetworks for learning Stochastic dynamical systems, algorithms for fine-tuning massive language models, and applications to Finance.
- 2022-2024 M.Sc. Joint: Giulia Livieri and Luca Galimberti, Alessio Spagnoletti (SNS Pisa), Enhancing time-series models with neural martingales

 Bachelors Student(s)
- 2024-2025 **Honors Thesis**, *Treavor Tidy*, McMaster University, Connections between proof theory and deep Learning, understanding how machines reason.
 - 2024 **Summer Research Project**, *Kyle Sung Joint with: Noah Forman*, McMaster University, Studying control of the Lipschitz constant of neural networks under descent algorithms, and its implications for extremal generalization bounds. Kyle completed a publishable paper (forthcoming).
 - 2023 **Summer Research Project**, *Ruiyang Hong*, McMaster University, Helped me with deriving upper bounds capacity bounds for hyperbolic neural network embeddings and verifying lower bounds for MLP embeddings for latent graph data. The paper was published in the top deep learning journal (Neural Networks).

Teaching Experience (McMaster University Only)

- 2025 (W) Computational Finance II, MFM 713, Graduate
- 2025 (W) Introduction to Analysis, MATH 3IA3
- 2024 (F) Calculus For Science I, MATH 1A03-C01
- 2024 (F) Computational Finance I (Lecture), MFM 703, Graduate
- 2024 W Computational Finance II, MFM 713, Graduate
- 2024 W Calculus For Science II, MATH 1AA3-C01
- 2023 F Calculus For Science I, MATH 1A03-C01
- 2023 F Computational Finance, MFM 703, Graduate
- 2023 W Calculus For Science II. MATH 1AA3-C01
- 2022 F Calculus For Science I. MATH 1A03-C01

Invited and Contributed Talks

- 2025 **TBD: Deep Learning in Finance**, *University of Oxford* Invited by: Blanka Hovarth
- 2025 **TBD: Deep Learning in Finance**, *London School of Economics* Invited by: Giulia Liviery
- 2024 Exponential Approximation Rates for Neural Operator Approximations to the Solution Operator of Certain FBSDEs, Invitation Carnegie Mellon University (CMU) Probability and Mathematical Finance Seminar Invited by: Johaness Wiesel
- Neural Operators Can Play Dynamic Stackelberg Games, Modeling, Learning and Understanding: Modern Challenges between Financial Mathematics, Financial Technology and Financial Economics Banff International Research Station (BIRS) Invited by: Antonis Papapantoleon
- 2024 Exponential Approximation Rates for Neural Operator Approximations to the Solution Operator of Certain FBSDEs, Eastern Conference on Mathematical Finance
 - Invited by: Silvana Pesenti
- 2024 An Approximation Theory for Metric Space-Valued Functions: From Rough Path Theory to Adapted Optimal Transport, 12th Bachelier World Congress in Mathematical Finance, FGV EMAp, Rio de Janeiro
- 2024 **Neural Operators Can Play Dynamic Stackelberg Games**, *CAIMS Annual Meeting*, Queens University, Kingston
- 2024 Byrne Conference on Stochastic Analysis in Finance and Insurance, *University* of Michigan, Ann Arbor Invited by: Ibrahim Ekren
- 2024 Pathwise Generalization bounds for Transformers, CMS Summer Meeting Session in: Mathematics of Machine Learning, University of Saskatoon, Saskatoon Invited by: Simone Brugiapaglia
- 2024 An Approximation Theory for Metric Space-Valued Functions With A View Towards Deep Learning, *DataSig Seminar Series*, University of Oxford, Oxford Invited by: Terry Lyons
- 2024 Universal Geometric Deep Learning via Geometric Attention, ml-ds-seminar, University of Vienna, Vienna Invited by: Philipp Petersen
- 2024 Universal Geometric Deep Learning via Geometric Attention, Actuarial Science and Financial Mathematics seminar series, University of Waterloo, Waterloo Invited by: Mario Ghossoub
- 2024 An Approximation Theory for Metric Space-Valued Functions With A View Towards Deep Learning, Applied Mathematics Seminars Centre de Recherches Mathématiques (CRM), McGill University, Montreal Invited by: Simone Brugiapaglia

- 2024 Some Recent Advances in Geometric Deep Learning in Finance, MAFI Seminar, ETH Zürich, Zürich Invited by: Josef Teichmann
- 2023 An Approximation Theory for Metric Space-Valued Functions With A View Towards Deep Learning, GAML Second Workshop on Geometry and Machine Learning, Max Plank, Leipzig Invited by: Anna Wienhard
- 2023 Fields-CFI Conference on Recent Advances in Mathematical Finance and Insurance, Stable Dimension Reduction and Efficient Conditional Distribution Approximation for Volterra Processes, Fields, Toronto
- 2023 The Transfer Principal: Universal Approximators Between Metric Spaces From Euclidean Universal Approximators, Bachelier Colloquium 2023, Métabief, France
- 2022 Measure-Valued Universal Approximation, Finance Al Seminar CIBC, Risk Management Group, Toronto, Canada Invited by: Hany Farag
- 2022 **Generic Geometric Priors via Probabilistic Transformers**, *Rice University's: MATH+X Symposium*, Conservatorio Castella, Costra Rica Invited by: Maarten de Hoop
- 2022 Causal Universal Approximation The Power to Approximate Structured Time-Series, Information Geometry Seminar, Stony Brook University, Remote Invited by: Pawel Polak
- 2022 **Universal Measure-Valued Neural Networks**, *Quantitative Finance Seminar*, BMO Capital Markets, Model Risk Group, Toronto, Canada
- 2022 Embedding Guarantees for Representations by Small Probabilistic Graph Transformers, GAML First Workshop on Geometry and Machine Learning, Heidelberg Institute for Theoretical Studies, Heidelberg Invited by: Anna Wienhard
- 2022 **Regular Conditional Distributions via Probabilistic Transformers**, *Bachelier Colloquium*, Metabief, France
- 2021 Universal Approximation with Exact Constraint Satisfaction is Possible with Transformers, *TU Munich, Munich, Munich, Germany*
- 2021 Geometric Deep Learning for High-Dimensional Option Pricing and Constrained Approximation, AIMS Lab Seminar, McMaster University, Hamilton, Canada
- 2021 Universal Approximation with Exact Constraint Satisfaction is Possible with Transformers, Talks in Financial and Insurance Mathematics, ETH Zürich Invited by: Josef Teichmann
- 2021 Universal Probability Measure-Valued Deep Neural Networks, *Probability Days*, University of Mannheim, Remote
- 2019 **Universal Approximation Theorems**, *Vienna Congress on Mathematical Finance VCMF 2019*, Vienna, Austria

- 2019 **Arbitrage-Free Regularization**, 10th Freiburg-Wien-Zürich (FWZ) Seminar, Vienne, Autstria
- 2019 Risk-Averse Conditional Expectation and Shortfall-Regression, ETH Zürich, Zürich, Switzerland
- 2018 Geometric Learning and Filtering in Finance, 10^{th} Bachelier World Congres, Dublin, Ireland
- 2018 Arbitrage-Free Regularization, 10^{th} Bachelier World Congres, Dublin, Ireland
- 2018 **Arbitrage-Free Regularization**, 5^{th} Workshop on Insurance Mathematics with a Special Session on Big-Data and Machine Learning in Risk Management, Montreal, Canada
- 2017 A Geometric Approach to Arbitrage-Free Modeling, Estimation, and Prediction, International Conference on Mathematical Finance, Waterloo, Canada Presentation by coauthor.
- 2017 **Geometric Learning and Non-Euclidean Filtering in Finance**, Canadian Mathematical Society Winter Meeting, Waterloo, Canada
- 2017 A Geometric Approach to Arbitrage-Free Modeling, Estimation, and Prediction, 11^{th} Bachelier Colloquium on Mathematical Finance and Stochastic Calculus, Métabief, France Presentation by coauthor.
- 2016 Manifold Learning Algorithms for Arbitrage-Free Low-Dimensional on-Linear Model Selection, Pacific Institute for the Mathematical Sciences (PIMS); High-Frequency Trading Workshop, Edmonton, Canada
- 2016 A Geometric Approach to Arbitrage-Free Modeling, Estimation, and Prediction, Canadian Mathematical Society Winter Meeting, Niagra, Canada Presentation by coauthor.
- 2016 Manifold learning algorithms for arbitrage-free low-dimensional nonlinear model selection, AMS AMS Fall Meeting, Brunswick, USA
- 2016 **Arbitrage-Free Regularization for Forward Rates**, *ENCS Data Science Research Centre*, Montreal, Canada

Workshops and Online Seminars Multi-University Seminars

- 2024- **(Toronto) Deep Learning Theory Seminar**, *Founder*, Joint: The Vector Institute Local Seminar Coordination
- 2022- **PhiMac (Financial Mathematics) Seminar**, *McMaster University*, Hamilton, Canada
- 2022-2023 **PDE/Analysis Seminar**, *McMaster University*, Hamilton, Canada Co-organizer Main: Lia Bronsard
- 2021-2022 SADA Lab, Universität Basel, Basel, Switzerland
- 2018-2021 Talks in Financial and Insurance Mathematics, ETH Zürich, Zürich, Switzerland

University and Department Service

Committees and Organization

- 2022-2025 Equity Diversity and Inclusion Committee (Non-Binary Representation)
- 2022-2024 Colloquium Committee
- 2022-2025 Masters in Financial Mathematics (MFM) Co-Coordination and Events

Committees and Organization

- (1) Ph.D. Student Committees Mathematics
- (1) Ph.D. Student Committees Statistics
- (1) Ph.D. Student Committees Computational Science and Engineering
- (1) MS.c. Student Committees Mathematics
- (1) MS.c. Student Committees Computational Science and Engineering

Reviewing Activities

- (4) The Journal of Machine Learning Research
- (3) Neural Networks
- (10+) Neural Information Processing Systems (NeurIPS)
- (10+) International Conference on Learning Representations (ICLR)
 - (3) Artificial Intelligence and Statistics (AISTATS)
 - (1) SIAM Journal on Financial Mathematics
 - (1) SIAM Journal on Applied Mathematics
 - (1) Physica D: Nonlinear Phenomena
 - (1) Journal of Optimization Theory and applications
 - (1) Sampling Theory, Signal Processing, and Data Analysis
 - (1) Annals of Mathematics and Artificial Intelligence

References

Josef Teichmann, ETH Zürich, josef.teichmann@math.ethz.ch Beatrice Acciaio, ETH Zürich, beatrice.acciaio@math.ethz.ch Juan-Pablo Ortega, NTU, juan-pablo.ortega@ntu.edu.sg Maarten do Hoop, RICE University, mvd2@rice.edu Nils Detering, Universität Düsseldorf:, Nils.Detering@hhu.de

Languages

Native, English and French

Fluent. Greek

Pre-Intermediate, Farsi

Beginner, German

Citizenship: Canadian