Theoretical foundations of learning robust neural ODEs

1 Supervisors

The thesis will be done in collaboration with Informatics Laboratory, HUN-REN Institute for Computer Science and Control (SZTAKI), Hungary. SZTAKI is expected to provide 50% of the funding and the student is expected to pass half of its time at SZTAKI.

The main advisors (directeur de thèse) will be Dr. Mihaly Petreczky (HDR, CRN CNRS) and Andás Benczúr (senior researcher, head of Informatics Laboratory, SZTAKI). the second advisors (co-encadrant) will be Dr. Ying Tang (maître de conférences, Université de Lille) from research team SHOC of the thematic group CO2 of the research laboratory CRIStAL (UMR CNRS 9189), and Dr. Bálint Daróczy (permanent researchers, SZTAKI).

Short CV of the supervisors

Mihaly Petreczky received the Ph.D. degree from Vrije Universiteit in Amsterdam, The Netherlands in 2006 and the HDR degree from Université Lille in 2023. In the past, he was a postdoc at Johns Hopkins University, USA (2006 - 2007), Eindhoven University of Technology, The Netherlands (2007-2009) and assistant professor at Maastricht University, The Netherlands (2009 – 2011) and at Ecole des Mines de Douai, France (2011 - 2015). He is currently a CNRS researcher at Centre de Recherche en Informatique, Signal et Automatique de Lille (CRIStAL), UMR CNRS 9189, France. His research interests include data-driven modelling of cyber-physical systems for control and its synergy with machine learning and statistics. In particular, he is interested in theoretically sound algorithms for learning models of cyber-physical systems and simplifying existing models. He co-authored 26 journal papers, some of them in high-impact journals, IEEE Trans. Automatic Control (7), Automatica (9), Systems & Control Lett. (4), SIAM J. Control (1), ESAIM COCV (2), Int. Journal of Nonlinear and Robust Control (2), NAHS (2), 3 book chapters, over 40 peer reviewed conference papers, leading conferences AAAI, CDC, ACC, ECC, IFAC World Congress, HSCC, ADHS, MTNS,NOLCOS. He is the scientific coordinator of CIFRE 'Motion dynamics modeling for fall of humans or two-wheeled vehicles (motorcycles/bicycles) : Combining domain knowledge-based and data driven models', Autoliv, and he was the scientific coordinator of the research contract 'Reliable AI for cyber-physical systems using control theory', IRT System-X, CNRS IEA 'Stability of learning algorithms for deep and recurrent neural networks by using geometry and control theory via understanding the role of overparameterization', the regional project CPER Data 'Machine learning meets Control', CNRS PEPS Blanc 2019, "PAC-Bayesian theory for recurrent neural networks: a control theoretic approach', regional project 'Estimation distribuée de systèmes dynamiques en réseaux' 2013-2017. He is an active member of the GDR MACS action on IA and Control, and an associate editor of Systems & Control Letters.

András Benczúr received his Ph.D. in 1997 in applied mathematics from the Massachusetts Institute of Technology. At present, he is senior researcher at the Institute for Computer Science and Control, Hungarian Research Network, and the scientific director of the Artificial Intelligence National Laboratory Hungary, a consortium of 11 institutions and over 200 researchers. His research revolves primarily around data mining, machine learning and web search, and he has represented his current institution as a principal investigator for multiple European Union and national R&D projects.

Dr.Ying Tang She received the M.S. Degree in Systems and Control Theory from the Institut Polytechnique de Grenoble, France, in 2012. She received the Ph.D. in Automatic Control from Grenoble Alpes University (Gipsa-lab), France in 2015. From 2015 to 2017 she was post-doc at CRAN, Nancy, France. Since September 2017 she is Assistant Professor at Lille University. Her research interests are in stability, analysis and parameter estimation of nonlinear systems with applications to traffic control problems and neural models.

Dr. Bálint Daróczy From December 2007 he worked on industrial and research projects related to machine learning, visual and text processing and multimodal search engines at the Institute for Computer Science and Control (SZTAKI), Eötvös Loránd Research Network (formerly part of the Hungarian Academy of Sciences) in Budapest. From 2010 he started to teach data mining and machine learning related courses at the Budapest

University of Technology and Economics (BME). Beside teaching he was supervising 8 BSc and 10 MSc theses and he currently co-supervises a PhD student at the Mathematical doctoral school at Eötvös Loránd University. He defended my PhD thesis at Eötvös Loránd University, Budapest, Hungary (title: "Machine learning methods for multimedia information retrieval", supervisor András Benczúr, PhD) in February 2017 with summa cum laude. In 2018 he received MTA Premium Postdoctoral Grant from the Hungarian Academy of Sciences for his research project "Manifolds and deep structures" and as a continuation he worked as a postdoctoral researcher funded by the grant at SZTAKI. Between November 2020 and October 2022 he joined professor Julien Hendrickx group at INMA at Université catholique de Louvain, Louvain-la-Neuve, Belgium as a postdoctoral researcher in the MIS "Learning from Pairwise Comparisons" of the F.R.S.-FNRS project. From November 2022 he will be full time research fellow at SZTAKI and continue his research project.

2 Description of the research project

Neural Ordinary Differential Equations (neural ODEs) [25, 21, 3, 36] are dynamical models blending neural networks and dynamical systems. They have gained popularity in the recent years mainly because of their attractive properties, e.g., ability to take into account physical constraints, to model irregularly sampled time-series, provide classifiers which are robust to adversarial attacks, predict long range dependencies, ability to serve as generative models (images, etc.) Despite many advances, there are important gaps in learning theory of such systems. One such gap is lack of formal guarantees for the generalisation error, i.e., the consistency of the learning algorithm. In this thesis we aim at filling this gaps by formalising neural ODEs and providing theoretical guarantees for learning robust neural ODEs.

Scientific methodology

The methodology relies on combining ideas from control theory on stability and robustness of dynamical systems and their learning with approaches from statistical machine learning.

Learning dynamical systems from data is the subject of system identification [23], which is a subfield of control theory. System identification provides a rich literature on the theoretical properties of such learning algorithms. However, the existing literature focuses on classes of dynamical systems relevant for control theory, and hence cannot be applied directly to general neural ODEs. Combining knowledge from system identification and machine learning, we aim at solving the following research problems

- **Robustness** Intuitively, robustness of a model means that small perturbations in inputs (or the distribution of inputs) will not result in a significant change of the label predicted by the model. One may argue that non-robust models are of limited use for prediction. In this thesis we plan to concentrate on robust neural ODEs. The notion of robustness is central to control theory, and it has been extensively studied there. In particular, formalisations of various versions of robustness are known such as input-to-output stability, L_p gains, input convergence, etc. Furthermore, various computationally effective conditions for checking robustness exist. We plan to apply these notions to neural ODEs. The challenge is that most of the results are geared towards classes of dynamical systems which are used for designing controllers. The class of neural ODEs is more general than the popular classes of dynamical systems studied in control. Moreover, neural ODEs are used primarily for prediction and not for control. Hence, it will be necessary to extend and adapt these results to neural ODEs.

- **Provide theoretical guarantees for statistical consistency of learning robust neural ODEs.** That is, we would liked to show that if a large enough number of data points is used for learning, then the generalisation error of neural ODEs learned from these data will converge to some lower bound representing the intrinsic error of modeling the underlying phenomenon by neural ODEs. We plan to extend recent results on statistical consistency of system identification algorithms [27, 5] and the results on convergence of stochastic gradient descend [6, 29]

- **Provide analytic PAC and PAC-Bayesian error bounds [15, 26] for neural ODEs.** The goal is to relate the generalisation error (prediction error on unseen data) with the prediction error on the training data. We plan to extend existing work [10, 31, 30, 11].

State of the art Machine learning and control theory are two closely related subjects with common roots. Recently, the two topics started to converge again: control theorists are becoming increasingly interested in using

machine learning techniques, while researchers in machine learning start looking at control problems and at possibilities to use results from control theory for machine learning. Despite this synergy, there is little prior work on applying ideas from control theory for proving consistency and error bounds for neural ODEs.

Regarding robustness of neural ODEs, there were a few attempts to formalize the notion of robustness [37, 17, 4], however, these formalizations are not yet satisfactory and they do not propose computationally efficient parameterizations of robust neural ODEs. Over the last few years, PAC- and PAC-Bayesian analyses have been conducted to explain the generalization performance of deep neural networks [9, 28], but most of the studies are limited to feed-forward neural networks, which are not suited for modelling dynamical systems in general, and for neural ODEs in particular. There are some prior work on PAC bounds for some classes of neural ODEs [12, 24, 16, 32, 22], but the resulting bounds have several disadvantages: they apply only to independently sampled multiple time-series and the bounds tend to grow exponentially with time. These disadvantages make them difficult to use for several learning tasks of interest, in particular, for learning models of physical systems. Moreover, the cited work does not address the role of robustness in generalization power.

Relevant track record by the supervision team The supervision team has already published preliminary results on the topic [31, 10, 11, 13, 14, 30, 29]. in leading conferences in control and machine learning . In particular, Mihaly Petreczky and his collaborators have successfully made the first steps towards PAC-Bayesian error bounds for learning dynamical systems [31, 10, 11, 30], and he has worked on system identification [5, 27, 1] and neural ODEs [8, 13, 14]. András Benczúr is a senior researcher in the field of machine learning and network science [18, 2, 19, 29, 20]. Bálint Daróczy has an extensive experience in theoretical properties of statistical learning algorithms, neural networks, and their applications [7, 19, 2, 29]. Ying Tang has an extensive experience on stability analysis and estimation of various classes of dynamical systems, which contain relevant classes of neural ODEs [35, 34, 33]. The supervisory team has history of collaboration [30, 29, 20].

In addition, the team secured funding on topics related to the thesis: the CPER Data project 'Machine learning meets control' 2018-2020, the CNRS PEPS Blanc 2019 project 'PAC-Bayesian theory for recurrent neural networks, and to the recently granted project 'Reliable AI for cyber-physical systems using control theory', which is a joint project with IRT System-X, and which is financed by the industrial initiative 'confiance.ia', and CNRS IEA 'Stability of learning algorithms for deep and recurrent neural networks by using geometry and control theory via understanding the role of overparameterization'. This indicates that the supervision team will be able to ensure that the prospective student achieves the stated goals.

Mihaly Petreczky co-supervised several students in the past, and he has earned his HDR degree in January 2023. Both Ying Tang and Bálint Daróczy are actively involved in the co-supervision of PhD students.

Relevance for the region and for the lab

The topic of the PhD thesis belongs to machine learning/AI, which is a topic of high priority both for the region and for CRIStAL. In fact, CRIStAL is an active participant in the regional project CPER Cornelia, and the proposed topic fits well the work package WP1 'bases théoriques et scientifiques de l'IA' of CPER Cornelia. In addition, the PhD project involves international collaboration with SZTAKI, the most prestigious research institute in computer science and control in Hungary. SZTAKI also provides 50% of funding, hence the proposed project is both international and co-funded by a third party. International and co-funded projects enjoy a high priority with the region. The main advisor from CRIStAL, Mihaly Petreczky has recently earned his HDR degree (2023) and he is not a main advisor of any PhD student at this point. This project would allow him to make the first step towards independent supervision of PhD students, and to do so on a topic which is important for CRIStAL and the region, and with international partners who contribute financaly to the project.

Work program

The work will be organised in the following work packages.

Work packages

WP1: Bibliography and state-of-the-art (0 - 6 month) The prospective PhD student will be expected to study the relevant literature and prepare a summary of the state of the art of the subject.

WP2: Statistical consistency of learning neural ODEs (12 - 28 month) The goal of this work package is to prove statistical consistency of learning algorithms for neural ODEs. This work package is likely to use the preliminary results [27, 5]. It can also be viewed as a preliminary step to WP3.

WP3: PAC(PAC-Bayesian) error bounds for neural ODEs (12-28 month) The purpose of this work package is to develop PAC and PAC-Bayesian error bounds for learning neural ODEs. This package can use the existing results [31, 10, 11, 13, 14, 30, 29]. Note that the PAC(-Bayesian) error bounds can also be used to show statistical consistency of learning algorithms, hence helping WP2.

WP4: Writing up the thesis and preparing the defence (27-36 month) This work package is devoted to writing the thesis and preparing its defence.

WP5: Dissemination (12-36 month) The PhD student is expected to publish his/her results in leading journals and conference proceedings in control (IEEE Trans. Automatic Control, Automatica, CDC, ACC) and machine learning (AAAI, NIPS, ICML, ECML, J. Machine Learning Research, Neurocomputing) and to present these papers in the leading conferences of both control theory and machine learning.

References

[1] Z. Alkhoury, M. Petreczky, and G. Mercère. Identifiability of affine linear parameter-varying models. *Automatica*, 80:62 – 74, 2017.

[2] F Ayala-Gómez, Bálint Zoltán Daróczy, ifj Benczúr András, M Mathioudakis, and A Gionis. Global citation recommendation using knowledge graphs. *JOURNAL OF INTELLIGENT & FUZZY SYSTEMS*, 34(5):3089–3100, 2018.

[3] Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.

[4] Krzysztof M Choromanski, Jared Quincy Davis, Valerii Likhosherstov, Xingyou Song, Jean-Jacques Slotine, Jacob Varley, Honglak Lee, Adrian Weller, and Vikas Sindhwani. Ode to an ode. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 3338–3350. Curran Associates, Inc., 2020.

[5] P. Cox, R. Tóth, and M. Petreczky. Towards efficient maximum likelihood estimation of lpv-ss models. *Automatica*, 2018.

[6] Bálint Daróczy, Rita Aleksziev, and András Benczúr. Tangent space separability in feedforward neural networks. *arXiv preprint arXiv:1912.09306*, 2019.

[7] Bálint Daróczy. Gaussian perturbations in relu networks and the arrangement of activation regions. *Mathematics*, 10(7), 2022.

[8] T. Defourneau and M. Petreczky. Realization theory of recurrent neural networks and rational systems. In *58th IEEE Conference on Decision and Control*, 2019.

[9] Gintare Karolina Dziugaite and Daniel M. Roy. Entropy-SGD optimizes the prior of a PAC-Bayes bound: Generalization properties of entropy-SGD and data-dependent priors. In *International Conference on Machine Learning*, 2018.

[10] D. Eringis, J. Leth, Zh.-H. Tan, R. Wisniewski, A. Fakhrizadeh Esfahani, and Petreczky M. Pac-bayesian theory for stochastic lti systems. In *60th IEEE Conference on Decision and Control*, 2021.

[11] Deividas Eringis, John Leth, Zheng-Hua Tan, Rafael Wisniewski, and Mihaly Petreczky. PAC-Bayesian bounds for learning LTI-ss systems with input from empirical loss. In *Workshop Frontiers4LCD ICML 2023*, Honolulu, United States, July 2023. arXiv admin note: text overlap with arXiv:2212.14838.

[12] Adeline Fermanian, Pierre Marion, Jean-Philippe Vert, and Gérard Biau. Framing rnn as a kernel method: A neural ode approach. *Advances in Neural Information Processing Systems*, 34:3121–3134, 2021.

[13] Martin Gonzalez, Thibault Defourneau, Hatem Hajri, and Mihaly Petreczky. Realization theory of recurrent neural odes using polynomial system embeddings. *Systems & Control Letters*, 173:105468, 2023.

[14] Martin Gonzalez, Hatem Hajri, Loic Cantat, and Mihaly Petreczky. Noisy Learning for Neural ODEs Acts as a Robustness Locus Widening. In *Workshop on Principles of Distribution Shift (PODS-ICML)*, Baltimore, United States, July 2022.

[15] B. Guedj. A Primer on PAC-Bayesian Learning. *arXiv preprint arXiv:1901.05353*, 2019.

[16] Joshua Hanson, Maxim Raginsky, and Eduardo Sontag. Learning recurrent neural net models of nonlinear systems. In *Learning for Dynamics and Control*, pages 425–435. PMLR, 2021.

[17] Ivan Dario Jimenez Rodriguez, Aaron D Ames, and Yisong Yue. Lyanet: A lyapunov framework for training neural odes. *arXiv*, 2022.

[18] Domokos Miklós Kelen and ifj Benczúr András. A probabilistic perspective on nearest neighbor for implicit recommendation. *International Journal of Data Science and Analytics*, 16:217–235, 2023.

[19] Domokos Miklós Kelen, Bálint Zoltán Daróczy, F Ayala-Gómez, Anna Ország, and ifj Benczúr András. Session recommendation via recurrent neural networks over fisher embedding vectors. *SENSORS*, 19(16), 2019.

[20] M. Domokos Kelen, Mihaly Petreczky, Péter Kersch, and A. Benczúr, András. Theoretical evaluation of asymmetric shapley values for root-cause analysis. In *ICDM*, 2023.

[21] Patrick Kidger. On neural differential equations. *arXiv preprint arXiv:2202.02435*, 2022.

[22] Pirkko Kuusela, Daniel Ocone, and Eduardo D Sontag. Learning complexity dimensions for a continuous-time control system. *SIAM journal on control and optimization*, 43(3):872–898, 2004.

[23] L. Ljung. *System Identification: Theory for the user (2nd Ed.)*. PTR Prentice Hall., Upper Saddle River, USA, 1999.

[24] Pierre Marion. Generalization bounds for neural ordinary differential equations and deep residual networks. *arXiv preprint arXiv:2305.06648*, 2023.

[25] Stefano Massaroli, Michael Poli, Jinkyoo Park, Atsushi Yamashita, and Hajime Asama. Dissecting neural odes. *Advances in Neural Information Processing Systems*, 33:3952–3963, 2020.

[26] David McAllester. Some PAC-Bayesian theorems. *Machine Learning*, 37(3), 1999.

[27] M. Mejari and M. Petreczky. Realization and identification algorithm for stochastic LPV state-space models with exogenous inputs. In *3rd IFAC Workshop on Linear Parameter-Varying Systems*, Eindhoven, Netherlands, 2019.

[28] Behnam Neyshabur, Srinadh Bhojanapalli, and Nathan Srebro. A PAC-bayesian approach to spectrally-normalized margin bounds for neural networks. In *International Conference on Learning Representations*, 2018.

[29] Dániel Rácz, Mihály Petreczky, András Csertán, and Bálint Daróczy. Optimization dependent generalization bound for reLU networks based on sensitivity in the tangent bundle. In *OPT 2023: Optimization for Machine Learning*, 2023.

[30] Dániel Rácz, Mihály Petreczky, and Bálint Daróczy. Pac bounds of continuous linear parameter-varying systems related to neural odes, 2023.

[31] V. Shalaeva, A. Fakhrizadeh Esfahani, P. Germain, and M. Petreczky. Improved pac-bayesian bounds for linear regression. In *34th AAAI Conference on Artificial Intelligence*, 2020.

[32] Eduardo D Sontag. A learning result for continuous-time recurrent neural networks. *Systems & control letters*, 34(3):151–158, 1998.

[33] Ying Tang, Alessio Franci, and Romain Postoyan. On-line detection of qualitative dynamical changes in nonlinear systems: The resting-oscillation case. *Automatica*, 100:17–28, 2019.

[34] Ying Tang, Christophe Prieur, and Antoine Girard. Stability analysis of a singularly perturbed coupled ode-pde system. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 4591–4596, 2015.

[35] Xinyong Wang, Ying Tang, Christophe Fiter, and Laurentiu Hetel. Sampled-data distributed control for homo-directional linear hyperbolic system with spatially sampled state measurements. *Automatica*, 139:110183, 2022.

[36] Ee Weinan. A proposal on machine learning via dynamical systems. *Communications in Mathematics and Statistics*, 1(5):1–11, 2017.

[37] Muhammad Zakwan, Liang Xu, and Giancarlo Ferrari-Trecate. On robust classification using contractive hamiltonian neural odes. *arXiv preprint arXiv:2203.11805*, 2022.

SZTAKI

November 30, 2023

Dr. András A. Benczúr

Institute for Computer Science and Control

Hungarian Research Network

13-17 Kende utca

H-1111 Budapest, Hungary

To whom it may concern,

I, undersigned, Dr. András Benczúr, the head of the Informatics Laboratory of Institute for Computer Science and Control, (SZTAKI), Hungarian Research Network, confirm that our organization, in collaboration with other Hungarian partners, is ready to fund 50% of a PhD scholarship for the collaborative project

> Theoretical foundations of learning robust neural ODEs proposed by Dr. Bálint Daróczy (SZTAKI), Dr. Mihály Petreczky and Dr. Ying Tang (UMR 9189 CRIStAL, Lile, France).

This contribution is conditional CRIStAL contributing the other 50% of the funding and the PhD student spending 50% of his/her time at SZTAKI (at least 18 months in total).

Please do not hesitate to contact me, if you need any additional information,

Sincerely,

Számítástechnikai és
Automatizálási Kutatóintézet
1111 Budapest XI., Kende u. 13-17.
MÁK 10032000-01738588-00000000
Adószám: 15300399-2-43 • 5.

Dr. András Benczúr