Final Scientific Report of the NKFIH KH_17 Project 125 698

MARKOV DECISION PROCESSES: ESTIMATION AND APPROXIMATION METHODS

(from September 1, 2017 till August 31, 2019)

1. Participants of the Project

- Dr. Balázs Csanád Csáji (Principal Investigator, Senior Research Fellow at MTA SZTAKI)
- Júlia Bergmann (initially: M.Sc. Student in Mathematics at BME, employed at MTA SZTAKI)
- Ambrus Tamás (M.Sc. Student in Applied Mathematics at ELTE, employed at MTA SZTAKI)
- Krisztián Balázs Kis (Software Developer at MTA SZTAKI)

2. Scientific Results and Related Activities

The main aim of the project was to make foundational studies which help dealing with estimation and approximation problems arising in *Markov decision processes* (MDPs). The two main research directions of the project, as described in the research plan of the project proposal, were as follows. The first one was motivated by the fundamental *exploration-exploitation* dilemma of MDPs and focused on constructing *non-asymptotic confidence regions*, with a special emphasis on dynamical systems. The second direction aimed at analyzing randomized methods in estimation and control problems, where *stochastic approximation* algorithms were highlighted, as they are especially important for model-free solutions of MDPs.

2.1 Results about non-asymptotic confidence region constructions

A new correlation-type, randomized, non-asymptotic, distribution-free *confidence region* construction, called Sign-Perturbed Correlation Regions (SPCR), was suggested [P3][P11], in collaboration with A. Carè, M. C. Campi and E. Weyer. SPCR combines the computational advantages of the LSCR (Leave-out Sign-dominant Correlation Regions) method with the *exact confidence* of the SPS (Sign-Perturbed Sums) algorithm. The ideas were demonstrated on ARX (autoregressive exogenous) type dynamical systems.

The potential improvements of (randomized) Data Perturbation (DP) methods with *input design* to minimize the expected volumes of the resulting (exact, non-asymptotic, distribution-free) confidence regions w.r.t. given input constraints, was analyzed for the case of linear regression problems [P10], in collaboration with S. Kolumbán. A Monte Carlo based *randomized* algorithm was suggested as an approximate solution and an explicit formulation of the fundamental ellipsoids building up DP confidence sets, in terms of the regressor and perturbation matrices, the true parameter and the noise, were also given [P10].

The problem of incorporating various *regularization* techniques (from Tikhonov regularization and LASSO to elastic nets regularization) in finite-sample confidence region constructions was addressed at the mini-symposium of ECMI and by a related book chapter [P5][P12]. These results led to the extension of SPS to general kernel methods [P1], in collaboration with K. B. Kis. A general data-driven framework was suggested to build non-asymptotic, distribution-free confidence regions around models constructed by kernel methods. It was proved that the regions have *exact* coverage probabilities for the *ideal* representation of the true underlying function, where ``ideal'' representations are functions from a Reproducing Kernel Hilbert Space (RKHS) with the property that their outputs are the same as the (noise-free) outputs of the underlying true functions, for all available inputs. The method was also demonstrated via several examples

and numerical experiments, e.g., using LS-SVM (least-squares support vector machine) classification, KRR (kernel ridge regression), SVR (support vector regression) and kLASSO (the kernelized version of LASSO).

Binary *classification* problems were also studied, in collaboration with A. Tamás, with the aim of estimating the underlying regression function which is the conditional expectation of the class labels given the inputs [P7]. The regression function is the key component of the *Bayes optimal classifier*, moreover, besides providing optimal predictions, it can also assess the risk of misclassification. Three *kernel-based* semi-parametric *resampling* methods were suggested to build non-asymptotic confidence regions for the regression function. We have proven that all of them guarantee confidence regions with *exact coverage* probabilities and they are *strongly consistent*. The results were also demonstrated on simulation studies [P7].

A survey (in Hungarian) of the theoretical results related to the (exact, non-asymptotic, distribution-free) confidence region construction of the SPS method, for stochastic linear systems, was also given [P4].

2.2 Results about stochastic approximation methods

Regarding stochastic approximation algorithms, first the effects of the *momentum* (heavy-ball) acceleration method on the classical LMS (Least Mean Square) adaptive filtering algorithm was investigated [P9], in collaboration with L. Gerencsér and S. Sabanis, where *weak convergence* results were obtained, under the assumption of stationary, ergodic and mixing signals, for both the standard and the accelerated variants. These results were then applied to express the asymptotic *covariance matrix* of momentum LMS, in terms that of the standard LMS and the gains of the correction and momentum terms.

Then, a key mathematical technology within the theory of stochastic approximation in a Markovian framework was revisited, in collaboration with A. Carè, B. Gerencsér, L. Gerencsér and M. Rásonyi. The *existence*, *uniqueness* and *Lipschitz continuit*y of the solutions of parameter-dependent Poisson equations were studied based on a new stability theory for Markov processes developed by Hairer and Mattingly. A transparent analysis of parameter-dependent Poisson equations with convenient conditions was provided that can be a vital tool for the ODE analysis of stochastic approximation in a Markovian framework [P6].

2.3 Industrial and engineering applications

An application paper [P2], which is a result of a collaboration between GE and MTA SZTAKI, describes results of an industrial (smart city) project related to an *analytical module* that processes the signals of a *wireless multi-sensor network* (WSN). *Confidence region* constructions were applied to provide prediction regions and reliability estimates for the analytics (namely, short-term forecasts and smoothed maps).

As a result of a collaboration between MTA SZTAKI, WZL RWTH Aachen, Germany, and FhG IPT, Aachen, Germany, stochastic approximation methods were applied to enhance the performance of an automated *stem cell* production platform, which approach led to up to 30% increase in the expected throughput [P8].

2.4 Relevant educational activities during the reported period

- B. Cs. Csáji gave *Ph.D. courses* for computer science students called "*Stochastic Models and Adaptive Algorithms*" at the Ph.D. School of Computer Science, Eötvös Loránd University (ELTE), Budapest, Hungary (Spring Semester, 2017/18), and at the Ph.D. School of Computer Science, Budapest University of Technology and Economics (BME), Budapest, Hungary (Autumn Semester, 2017/18).
- B. Cs. Csáji gave an *M.Sc. course* for mathematician students called "*Markov Decision Processes and Reinforcement Learning*" (Spring Semester, 2018/19), Institute of Mathematics, Faculty of Natural Sciences, Budapest University of Technology and Economics (BME), Budapest, Hungary.

- B. Cs. Csáji has supervised several (M.Sc. and B.Sc.) theses related to the project, for example, estimating stochastic linear systems with changing dynamics was studied [T1]; the extension of SPS for the least absolute deviation (LAD) estimate was investigated [T2]; the effects of momentum acceleration on the Q-learning algorithm was examined [T3]; stochastic approximation methods working in MDPs with changing-dynamics was reviewed; [T4] as well as stochastic multi-armed bandits were studied [T5].

2.5 Scientific organizational activities and achievements during the project

- B. Cs. Csáji has *organized* and *chaired*, together with A. Carè (University of Brescia, Italy), the mini-symposium "*Finite-Sample System Identification*" at the 20th European Conference on Mathematics for Industry (ECMI), Budapest, Hungary, 18-22 June, 2018.
- B. Cs. Csáji was a member of the *International Program Committee* of the 13th IFAC ALCOS (Workshop on Adaptive and Learning Control Systems), Guildhall Winchester, Winchester, UK, 2019.
- B. Cs. Csáji was a member of the *International Program Committee* of the 15th IFAC Symposium on Large Scale Complex Systems, Delft, The Netherlands, May 26-28, 2019.
- B. Cs. Csáji has *chaired* the session "*Input Design*" at the 18th IFAC Symposium on System Identification (SYSID 2018), Stockholm, Sweden, July 9-11, 2018.
- B. Cs. Csáji has *chaired* the session "*Estimation IV*" at the 57th IEEE Conference on Decision and Control (CDC 2018), Miami Beach, Florida, December 17-19, 2018.
- B. Cs. Csáji gave a *Plenary Talk* with the title "*Statistical Learning Theory: Classification and Regression with Stochastic Guarantees*" at the 33rd Hungarian Conference on Operational Research, Szeged, 2019.
- B. Cs. Csáji has received a *Publication Award* from MTA SZTAKI: The Institute for Computer Science and Control, Hungarian Academy of Sciences, Budapest, Hungary, 2018.
- J. Bergmann has won the 1st prize (Applied Pedagogy) and a Pro Progressio Special Award at the *Scientific Student Conference* (TDK) of the Budapest University of Technology and Economics, 2018.
- A. Tamás, a student participant, has won a *Regular Research Scholarship* (based on his results during the project) from ELTE-TTK: Eötvös Loránd University, Faculty of Science, Budapest, Hungary, 2019.
- B. Cs. Csáji was a *Guest Editor*, together with L. Gerencsér, of a special issue of the "Alkalmazott Matematikai Lapok" (AML) journal. He also became a regular member of the *Editorial Board* of AML.
- B. Cs. Csáji has been a member of the International Federation of Automatic Control (IFAC) committees:
  - IFAC Technical Committee (1.1): *Modelling, Identification and Signal Processing*
  - IFAC Technical Committee (1.2): *Adaptive and Learning Systems*
  - IFAC Technical Committee (5.4): *Large Scale Complex Systems*
- B. Cs. Csáji was a *reviewer* for the following prestigious international journals (during the two years of the project): *Automatica*; *IEEE Transactions on Automatic Control*; and *IEEE Control Systems Letters*.
- B. Cs. Csáji was a *reviewer* for the following top international conferences (during the project):
  - AISTATS 2018 and 2019: *International Conference on Artificial Intelligence and Statistics*;
  - ICML 2018 and 2019: *International Conference on Machine Learning*;
  - IJCAI 2019: *International Joint Conference on Artificial Intelligence*;
  - NIPS 2018 and 2019: *Neural Information Processing Systems*;
  - UAI 2019: *Uncertainty in Artificial Intelligence;*
  - IEEE CDC 2018 and 2019: *Conference on Decision and Control*;
  - IFAC SYSID 2018: *Symposium on System Identification*;
  - ECC 2018: *European Control Conference*;
  - IEEE CCTA 2019: *Conference on Control Technology and Applications*.

3. Publications, Presentations and Supervised Theses

3.1 Published or accepted (peer-reviewed) journal papers during the project

[P1] Csáji, B. Cs.; Kis, K. B.: "*Distribution-Free Uncertainty Quantification for Kernel Methods by Gradient Perturbations*", Machine Learning, Springer, Vol. 108, 2019, pp. 1677–1699 (Impact Factor: 2.809)

[P2] Csáji, B. Cs.; Kemény, Zs.; Pedone, G.; Kuti, A.; Váncza, J.: "*Wireless Multi-Sensor Networks for Smart Cities: A Prototype System with Statistical Data Analysis*", IEEE Sensors Journal, Vol. 17, Issue 23, Dec 1, 2017, pp. 7667–7676 (Impact Factor: 2.617)

[P3] Carè, A.; Csáji, B. Cs.; Campi, M. C.; Weyer, E.: "*Finite-Sample System Identification: An Overview and a New Correlation Method*", IEEE Control Systems Letters, 2018, pp. 61–66

[P4] Csáji, B. Cs.: "*Szimmetria és konfidencia*" (Symmetry and Confidence), "Alkalmazott Matematikai Lapok" (Hungarian Applied Mathematical Journal), Volume 36, Issue 2, 2019 (in print)

Published or accepted (peer-reviewed) book chapters during the project

[P5] Csáji, B. Cs.: "*Non-Asymptotic Confidence Regions for Regularized Linear Regression Estimates*", Progress in Industrial Mathematics at ECMI, Mathematics in Industry, Springer, 2019 (in print)

3.2 Published or accepted (peer-reviewed) conference papers during the project

[P6] Carè; A.; Csáji, B. Cs.; Gerencsér, B.; Gerencsér, L; Rásonyi, M.: "*Parameter-Dependent Poisson Equations: Tools for Stochastic Approximation in a Markovian Framework*", 58th IEEE Conference on Decision and Control (CDC), Nice, France, 2019 (in print)

[P7] Csáji, B. Cs., Tamás, A.: "*Semi-Parametric Uncertainty Bounds for Binary Classification*", 58th IEEE Conference on Decision and Control (CDC), Nice, France, 2019 (in print)

[P8] Egri, P.; Csáji, B. Cs.; Kis, K. B.; Monostori, L.; Váncza, J.; Ochs, J.; Jung, S.; König, N.; Schmitt, R.; Brecher, C.; Pieske, S.; Wein, S.: "*Bio-Inspired Control of Automated Stem Cell Production*", 13th CIRP Conference on Intelligent Computation in Manufacturing Engineering, Naples, Italy, 2019 (accepted)

[P9] Gerencsér, L.; Csáji, B. Cs.; Sabanis, S.: "*Asymptotic Analysis of the LMS Algorithm with Momentum*", 57th IEEE Conference on Decision and Control (CDC), Miami, Florida, 2018, pp. 3062-3067

[P10] Kolumbán, S.; Csáji, B. Cs.: "*Towards D-Optimal Input Design for Finite-Sample System Identification*", 18th IFAC Symposium on System Identification (SYSID), Stockholm, Sweden, 2018, pp. 215-220

[P11] Carè, A.; Csáji, B. Cs.; Campi, M. C.; Erik, W.: "*Finite-Sample System Identification: An Overview and a New Correlation Method*", 56th IEEE Conference on Decision and Control (CDC), Melbourne, Australia, December 12-15, 2017, pp. 4612-4617 (conference version of journal paper [P3])

3.3 Other presentations with abstracts during the project

[P12] Csáji, B. Cs.: "*Regularization in Finite-Sample System Identification*" (abstract), The 20th European Conference on Mathematics for Industry (ECMI), Book of Abstracts, Budapest, Hungary, 2018

[P13] Carè, A.; Csáji, B. Cs.; Campi, M. C.; Erik, W.: "*Old and News Challenges in Finite-Sample System Identification*" (abstract), The 20th European Conference on Mathematics for Industry (ECMI), Budapest, Hungary, 18-22 June 2018, Book of Abstracts, pp. 248

[P14] Csáji, B. Cs., Tamás, A., Kis, K. B.: "*Non-Asymptotic Uncertainty Bounds for Kernel Estimates*" (abstract), 30th European Conference on Operational Research, June 23-26, Dublin, Ireland, 2019

[P15] Csáji, B. Cs.: "*Statisztikus tanuláselmélet: klasszifikáció és regresszió sztochasztikus garanciákkal*" (Statistical Learning Theory: Classification and Regression with Stochastic Guarantees, abstract), Plenary Talk at the 33rd Hungarian Conference on Operational Research (MOK), Szeged, 2019

3.4 Educational presentations and public lectures related to the project

[L1] B. Cs. Csáji gave a *public lecture* on "*Data and Analytics: Predictive Methods*" at the 1st INDIGO Industrial Digitization Day, Budapest, Hungary, October 10, 2017

[L2] B. Cs. Csáji gave a *public lecture* on "*In the Wake of Learning Machines*" at the 11th Digital Equal Opportunities (DE!) conference of the John von Neumann Computer Society (NJSZT), Danubius Hotel Gellért, Budapest, Hungary, November 28, 2017

[L3] B. Cs. Csáji gave a *presentation* on "*Non-Asymptotic Distribution-Free Statistical Inference for Regression Models*" at the workshop "Mathematical Foundations of Artificial Intelligence", jointly organized by MTA SZTAKI and the Alfréd Rényi Institute of Mathematics, Budapest, November 29, 2017

[L4] B. Cs. Csáji gave a *tutorial* on "*Stochastic Approximation*" at the Institute for Computer Science and Control, Hungarian Academy of Sciences (MTA SZTAKI), Budapest, Hungary, December 6, 2019

[L5] B. Cs. Csáji gave a *presentation* on "*Asymptotic Analysis of the LMS Algorithm with Momentum*" at the "Mathematical Foundations of Artificial Intelligence" Mini-Conference, Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Budapest, Hungary, April 24, 2019

3.5 Supervised master's and bachelor's theses related to the project

[T1] Bergmann, J.: "*Stochastic Linear Systems with Changing Dynamics*" (language: English), M.Sc. Thesis, Mathematics, Institute of Mathematics, Faculty of Natural Sciences, BME: Budapest University of Technology and Economics, Budapest, Hungary, 2018 (graduated)

[T2] Tamás, A.: "*Non-Asymptotic Confidence Regions for the Least Absolute Deviations Estimate*" (language: English), B.Sc. Thesis, Applied Mathematics, Institute of Mathematics, Faculty of Science, ELTE: Eötvös Loránd University, Budapest, Hungary, 2018 (graduated)

[T3] Juhász, M.: "*Momentum Acceleration for Q-learning*" (language: Hungarian), B.Sc. Thesis, Mathematics, Institute of Mathematics, Faculty of Natural Sciences, BME: Budapest University of Technology and Economics, Budapest, Hungary, 2019 (graduated)

[T4] Molnár, A. E.: "*Stochastic Approximation in Time-Varying Markov Decision Processes*" (language: Hungarian), B.Sc. Thesis, Mathematics, Institute of Mathematics, Faculty of Natural Sciences, BME: Budapest University of Technology and Economics, Budapest, Hungary, 2019 (graduated)

[T5] Kiss, B.: "*Bandit Algorithms in Stochastic Environments*" (language: Hungarian), B.Sc. Thesis, Mathematics, Institute of Mathematics, Faculty of Natural Sciences, BME: Budapest University of Technology and Economics, Budapest, Hungary, 2019 (graduated)

[T6] Füzesdi, M.: "*Stochastic Approximation and Reinforcement Learning*" (tentative title, language: English), M.Sc. Thesis, Mathematics, Institute of Mathematics, Faculty of Natural Sciences, BME: Budapest University of Technology and Economics, Budapest, Hungary, 2020 (under preparation)