Final Report for grant PD 121107

titled

Information dynamics of cooperative stochastic networks - - design and analysis

In this document we collect and summarize the research output during the grant period 2016.10.01.–2019.09.30. During quality exploratory mathematical research, there is the challenge coupled with the risk taken towards the goals outlined. This is reflected in the results achieved: some of them exactly match the problems specified in the proposal while some of them were adjusted on the way while pointing in the same direction of understanding.

The research carried out is not the end of the road, therefore we also summarize the ongoing efforts together with the perspectives for continuations.

Completed research

A surprising result obtained is on the topic of understanding mixing on variations of Small World Networks, in particular when only a cycle graph of n vertices is used with a low number of additional edges. It was found [3] that even a single edge can change the magnitude of the mixing time, if properly chosen and also a non-reversible Markov chain is utilized. On the cycle itself $\Theta(n^2)$ is the best possible mixing, this slight modification allows $\Theta(n^{3/2})$ to be achieved.

The preceding result on adding a complete graph on a sublinear, $n^\gamma, 0 < \gamma < 1$ number of random vertices [8] has gone through all necessary post-processing and is now accepted and published in print.

Another main topic of the grant was the analysis of the push-sum algorithm, designed for achieving consensus on a network when communication between nodes can be asynchronous. There are two natural questions: error of the consensus value from the true average, and rate of convergence to consensus. In the end we could get a strong result for the second one. The result applies for a general class of schemes, instead of iid updates it is sufficient to assume that the series of random matrices describing the updates form a stationary sequence. This includes having multiple nodes having complex updates in parallel, multi-step rules, etc. The result establishes a correspondence [5] between the convergence rate of this general push-sum scheme and the Lyapunov exponents of the series of random matrices, as classical random matrix products. Consequently, existing tools for spectral analysis of random matrix products can be carried over to characterize the convergence rate of push-sum algorithms.

The motivating result [9] on understanding push-sum with packet loss has gone through multiple iterations substantially improving quality and augmenting results, now in print.

A different aspect of consensus is its robustness to noisy measurements of the difference. An illustrative example is vehicle platoon formation, as mentioned in the grant proposal, where sensor measurements are noisy, collision avoidance is a priority, still distances should be kept low. By an affine transformation this is equivalent to the vehicles attempting to agree on an offset, while keeping their order. We succeeded in formulating an abstract result [12] that ensures convergence with the help of a Lyapunov function, relying on the fact that each agent is going in the right direction. This is sufficient to prohibit divergence or oscillations. The current proof is available for positive definite quadratic forms and selected positive semidefinite ones, and this is sufficient to be applicable for the platoon formation problem.

There has been considerable work on understanding factor of iid processes on infinite regular tree graphs. This is a fundamental building block that can be viewed from various aspects: as random variables at every vertex, they can be considered as a limit of random processes on graphs or network interactions as the size of the graph grows to infinity. When explored from one direction, it can be considered as a branching process with strong symmetry conditions imposed. The ergodic properties of such properties is a central question.

The decay of correlation between the value of two vertices as their distance grows has been well understood for scalar valued processes. However, correlation might now be the ultimate measure, even less so for discrete valued processes, where a remapping of the values would change the correlation without changing the actual dynamics. Therefore, we decided to work on understanding information theoretical structure of such processes. One of the papers focuses on local behavior [1], stating entropy inequalities, giving lower bounds on the entropy of the joint variables of nearby random variables in terms of the entropy of a single variable. Loosely speaking, these are quantitative statements of the form that nearby variables have to be “somewhat” independent.

The other part of this work considers long range comparison, on the mutual information decay of the random variables at far away nodes [7]. Interestingly, the answer changes depending on exactly how it is phrased. On the d-regular infinite tree there is a sharp global bound of the order of $(d-1)^{-k/2}$ for vertices at distance k. However, once a process is fixed, the rate of decay cannot be substantially better than $(d-1)^{-k}$ with the increase of k.

A different challenge for the extension of Markov chains is when the dynamics is time dependent. That is, a background stationary process is given, not necessarily Markovian, which determines the transition kernel at every time instance. The evolution of the main process is Markovian, but the transition kernel changes from time to time. This structure is not only interesting as an intermediate construct between arbitrary stochastic processes and Markov chains, but appears inherently in financial mathematical modeling. Our study [10] concerned proving convergence to a stationary distribution and also establishing an ergodic theorem (in spirit of a law of large numbers along the process). There are a number of conditions that the driving processes have to satisfy and then convergence, ergodicity follows, the optimization of these condition is still for future work.

A related issue appears in the presence of a feedback, when the dynamics is affected by the state of the process. This happens for instance in the case of the application of the ODE method for stochastic approximation, when a parametrized class of random variables is given together with an optimization problem. A possible approach is to iteratively update the parameter and evolve according to the corresponding dynamics to acquire (noisy) information about the value to be optimized. The current work [2] revisits the technical foundation of this approach, now putting it in a context so that the required conditions are easily verifiable.

Another element of understanding Markov chain mixing is for a Gibbs sampler, a special case of a Metropolis algorithm. Intuitively the question is how the misalignment of the possible steps and the density of the target measure slows down mixing, and how this can be made even worse by boundary effects. This abstract question can be modeled by having $[0,1]^2$ as the state space, allowing only horizontal or vertical moves, but targeting a distribution concentrated near the diagonal (proportional to $\exp(-A^2(x-y)^2)$ in particular). This problem was posed by Persi Diaconis, which also turns out to have a deep meaning for Bayesian statistics, for describing priors on the almost exchangeability of two datasets. The exact mixing order has been determined [4] for this Markov chain, confirming the conjectured $\Theta(A^2)$, improving on the previously known $O(\exp(A^2))$.

The converse view on ergodicity and the effect of noise is the analysis of controllability. A mature framework formulating this question is the theory of synchronizing automata. The goal is to apply a sequence of control steps to a system with only a limited choice available, driving any initial state to the same final state. Multiple questions arise, deciding whether this is possible, determining the minimal number of steps needed - the famous Černý’s conjecture - or analyzing the complexity of finding a (minimal) such synchronizing sequence. Our work [11] targets the class of automata where the control steps are rich enough to generate any mapping. We show that this is sufficient to confirm a quadratic length for the minimal synchronizing word, matching the order of growth conjectured in general. Additionally, a warning example is presented against a usual proof attempt of iterative synchronization, showing that a quadratic number of steps might be necessary just to bring one chosen pair of states to another one.

Finally, we add an admittedly off-topic result to the basket, in the area of Euclidean geometry. However, this outcome of branching curiosity ended up in solving a question open for more than 60 years, originally advertised by Erdős. What is the maximal number of points in $\mathbb{R}^d$ such that any angle determined by them is acute? By allowing right angles the upper bound 2^d has been shown since decades and demonstrated by the hypercube. With this further restriction, initially even a linear optimal growth was conjectured, refuted later to be exponential. The rate was still unknown and was improved at several steps. Finally, we established [6] that the optimal exponential rate is again 2, now only having a constant factor between the lower and upper bounds. A survey on the result will soon be published in the American Mathematical Monthly to reach a wider audience.

In terms of the deliverables set in the proposal, they have been met and the research turned out to be even more successful. Yearly two arXiv manuscripts were planned, with yearly two journal submissions in parallel. Summarizing the publications described above, during the three year period there were 10 arXiv uploads and 2 earlier manuscripts refined. Out of these 8 journal publications have been accepted with 3 journal submissions still under consideration, in most cases choosing high-quality Q1 rated journals. One manuscript is still is being extended from a conference exposition to a full journal version.

Additional scientific activity

Besides carrying out research, training was also planned during this period. It turned out that solitary study was less efficient, only partially going through the books as outlined. However, workshops and summer schools served as alternate opportunities to advance and to get up-to-date with current scientific progress. Right at start, there was the Markov chain Mixing Times workshop organized at the American Institute of Mathematics (San Diego, 2016). An event linking together areas was the Graph limits, groups and stochastic processes Summer School at the Rényi Institute (Budapest, 2017). At this moment, a Markov chain Monte Carlo school is coming up organized by the Hausdorff Center for Mathematics (Bonn, 2020).

During the period substantial effort has been put into building and maintaining the research network. This includes long standing collaborations with colleagues at Université catholique de Louvain (Belgium), fresh collaborations with University of Oregon (US) and new connections at Princeton University (US) and at Stanford University (US).

Perspectives

Several threads of the research project offer open questions and provide inspiration for future work. Some of them are already being investigated and are the target of various collaborations.

Jointly with David Levin we work on the model related to Small World Networks, investigating the case of increasing the probability of extra edges to ε/n. A number of previously comfortable independence claims break, challenging to develop a very fine-tuned understanding to result in mixing time bounds.

Together with Julien M. Hendrickx a similar work is about to conclude where the vertices on the cycle affected is polynomial, $n^\gamma, 0 < \gamma < 1$, but a sparse interconnection is used instead of a complete graph as in the preceding work. There is also an effort to generalize the result on robust consensus, to be able to handle more general Lyapunov functions in the framework built before.

In a work with Miklós Rásonyi, the analysis of Markov chains in random environment continues, with the goal of both relaxing the necessary conditions and also to formulate them in a way that they become easily verifiable.

In a team of five with Italian and Hungarian participants, stochastic approximation study is carried on, to exploit the newly found fundamental observations for the algorithmic constructions and obtain results in such implementations.

In view of the various research questions and the associated activity, funding is also being sought after. Therefore two grants have been submitted and are currently under consideration, one with the National Research, Development and Innovation Office - NKFIH and one with the European Research Council.

References

[1] Á. Backhausz, B. Gerencsér, and V. Harangi, *Entropy inequalities for factors of IID*, Groups, Geometry, and Dynamics, 13 (2019), pp. 389–414.

[2] A. Carè, B. C. Csáji, B. Gerencsér, L. Gerencsér, and M. Rásonyi, *On the Poisson equation of parameter-dependent Markov chains.* arXiv:1906.09464, 2019.

[3] B. Gerencsér, *Analysis of a non-reversible Markov chain speedup by a single edge.* arXiv:1905.03223, 2019.

[4] ——, *Mixing time of an unaligned Gibbs sampler on the square*, Stochastic Processes and their Applications, 129 (2019), pp. 3570–3584.

[5] B. Gerencsér and L. Gerencsér, *Tight bounds on the convergence rate of generalized ratio consensus algorithms.* arXiv:1901.11374, 2019.

[6] B. Gerencsér and V. Harangi, *Acute sets of exponentially optimal size*, Discrete & Computational Geometry, (2018). To appear.

[7] ——, *Mutual information decay for factors of i.i.d.*, Ergodic Theory and Dynamical Systems, 39 (2019), pp. 3015–3030.

[8] B. Gerencsér and J. M. Hendrickx, *Improved mixing rates of directed cycles by added connection*, Journal of Theoretical Probability, 32 (2019), pp. 684–701.

[9] ——, *Push-sum with transmission failures*, IEEE Transactions on Automatic Control, 64 (2019), pp. 1019–1033.

[10] B. Gerencsér and M. Rásonyi, *On the ergodicity of certain Markov chains in random environments.* arXiv:1807.03568, 2018.

[11] F. Gonze, V. V. Gusev, B. Gerencsér, R. M. Jungers, and M. V. Volkov, *On the interplay between Babai and Černý's conjectures*, in Developments in Language Theory, 2017, pp. 185–197.

[12] J. M. Hendrickx, B. Gerencsér, and B. Fidan, *Trajectory convergence from coordinate-wise decrease of quadratic energy functions, and applications to platoons*, IEEE Control Systems Letters, 4 (2019), pp. 151–156.