Final report: Structure, Chaos and Simplification, KKP133921

Balázs Szegedy

May 1, 2025

The main objective of the project was to explore methods that allow us to create simplified models from large, often chaotic structures by extracting the essence and filtering out noise—thereby leading to a better understanding of the structure. Among classical methods, this includes principal component analysis and Szemerédi's regularity method, along with various generalizations. Within the project, we further generalized these methods and connected them in novel ways.

The structures analyzed in the project can be described using graphs, Markov operators, time series, dynamical systems, and functions defined on groups. This motivates the focus of our theoretical research on higher-order Fourier analysis, ergodic theory, the theory of graph limits, and group theory.

However, the development of deep learning complements the project's topic with an intriguing perspective. In neural networks, simplified, low-dimensional models of large data sets appear, which can be regarded as far-reaching generalizations of principal component analysis. For this reason, we placed significant emphasis on incorporating a machine learning perspective as well. This also includes the question of how neural representations evolve during the learning process.

Results in higher-order Fourier analysis: Within the framework of the project, we placed significant emphasis on the development of higher-order Fourier analysis from both theoretical and practical perspectives. The main motivation for this is that classical Fourier analysis is a key method in both theoretical and applied mathematics, so we may hope that its higher-order generalization could also become a fundamentally important tool. We explored this area in great depth through a series of articles spanning nearly 400 pages [1, 2, 3, 4, 5, 6, 7, 8]. In these articles we reached the following goals.

- We further advanced nilspace theory which is the algebraic foundation of higher-order Fourier analysis. Using these results we proved inverse theorems for the Gowers norms in various families of Abelian groups[1, 2, 4, 6, 8].

- We analized the problem of arithmetic exchangeability and provided a full description of affine $\mathbb{F}_2^\omega$-exchangeable probability measures. Note that this is an extension of the famous De Finetti theorem known from statistics and probability theory [3].

- In the theory of dynamical systems, we described the structure of $\mathbb{F}_p^\omega$-systems of order k with methods in higher-order Foureire analysis [5].

- We further developed the nilspace approach to higher-order Foureir analysis (see Figure 1). We introduiced a new form of regularization using low dimensional nilspace representations and used it to prove a general arithmetic regularity theorem. Our publication appeared in the prestigious Crelles Journal [8].

- Based on all of the previous theoretical result, we developed efficient algorithms in higher-order Fourier analysis that can be used for decomposing functions into higher order harmonic functions, higher-order denoising of time series and time series prediction [4], see also Figure 2.

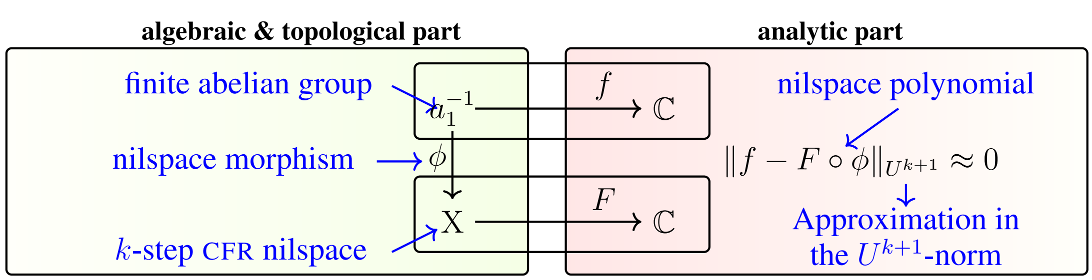

Figure 1: The nilspace approach to k-th order Fourier analysis

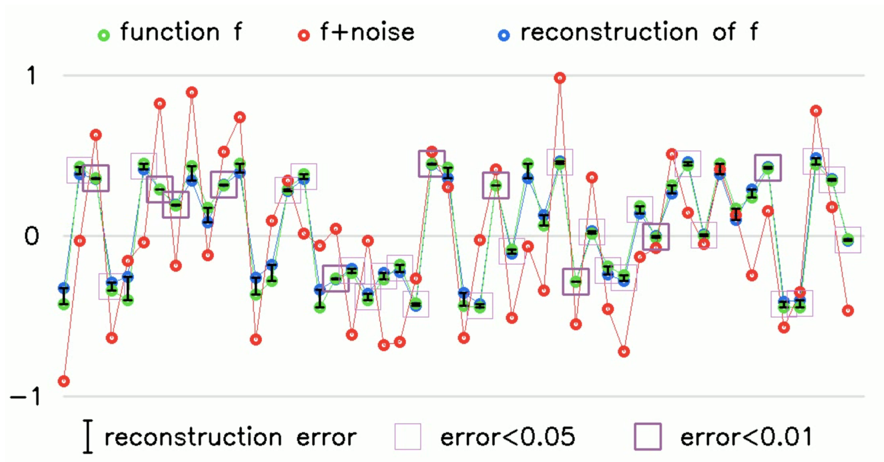

Figure 2: Removing added random noise from a quadratic structured function by the spectral algorithm on the cyclic group $\mathbb{Z}_{500}$. A window of length 50 is plotted for illustration. In this example, the function $f(i) := \sin(8i^2 + 3i + 1)$, $i \in [500]$ (green graph) is perturbed by random noise, resulting in the function $g = f + r$ (red graph). The spectral algorithm is applied to g and the reconstruction f_2 of f (blue graph) is obtained by the projection of g to the space spanned by the 6 leading eigenvectors of the operator constructed from g. The plot highlights the reconstruction error $|f(i) - f_2(i)|$.

Results in graph limit theory, probability theory and related fields: One of the biggest challenges in graph limit theory is to generalize results from the most complete limit theory—namely, dense graph limit theory—to sparse graph sequences. It was known before the start of the project that Markov operators can be used to represent sparse graph limits. However, sparse graph limit theory lacked subgraph densities, which are defined by Feynman-type integrals in the dense setting. A major breakthrough of the project was the generalization of subgraph densities to singular Markov kernels (satisfying certain smoothness conditions) representing sparse graph limits. Our publication [9] appeared in the top journal Advances in Mathematics. A closely related result [10] introduced natural probability distributions on orthogonal representations of graphs and calculated subgraph densities for special Markov kernels arising for high dimensional spheres. Note that introducing subgraph densities for sparse graph sequences opens up the intriguing possibility that Szemerédi's regularity lemma can be pushed into the sparse regime in a novel way. In this sense, our results motivate new research in dimension reduction of sparse structures.

Another important question in graph limit theory concerns random graphs. It is not know if random d-regular graphs converge in the stronger local-global topology (it is easy to see that they converge in Benjamini-Schramm topology). However it is known that this question is equivalent with describing certain automorphism invariant processes (called typical processes) on the infinite d-regular tree. This correspondence creates an interesting connection between random graph theory and dynamical systems. We made significant progress in this question by giving new entropy based conditions for typical processes in [11].

Results in deep learning: An important goal of the project was to lay the groundwork for a theoretical understanding of neural representations. In particular, useful low-dimensional representations of complex data emerge during the training process, making it natural to investigate training dynamics from a mathematical perspective. One of the main questions is to describe the evolution of the so-called neural tangent kernel. We conducted extensive computer experiments that revealed several intriguing phase transitions in the evolution of the quasidimension (a relaxed notion of rank) of the neural tangent kernel. These results are parts of an ongoing and yet unfinished research direction. In another related project we investigated neural tangent kernels in overparametrized learning situations [13]. We investigated the space of gradients coming from different momentum optimizers and developed a new adaptive optimization algorithm [14]. We conducted research related to the information theoretic aspects of machine learning and studid opptimal transport with f-divergence regularization [12].

Higher order theories: Motivatied by higher-order Fourier analysis, and classical representation theory of compact gropups, we introduced a novel higher order generalization of group theory [15] which is related to algebraic topology, category theory and potentially to mathematical physics. The development of this new theory is still in an early phase but there are promising sings for interesting applications.

Remarks: Some of the papers produced by the project have not yet been accepted. In certain cases it is due to the long referee process. In other cases it is due to risky journal and conference choices. Those papers will be sent to other journals and eventually published. There are also unfinished parts of the project, notably the research direction related to tensor networks. We accumulated manuscripts and partial results in this direction but publication is delayed.

Acknowledgement : The PI of the project is very grateful for the funding which was crucial for this research project.

References

[1] Candela, González-Sánchez, and Szegedy, *Free nilspaces, double-coset nilspaces, and Gowers norms.*, https://arxiv.org/abs/2305.11233, 2023.

[2] Candela, González-Sánchez, and Szegedy, *On higher-order Fourier analysis in characteristic p.* Ergodic Theory and Dynamical Systems, 43(12):3971–4040, Jan. 2023.

[3] Candela, González-Sánchez, and Szegedy, *On $\mathbb{F}_2^\omega$-affine-exchangeable probability measures* Ergodic Theory and Dynamical Systems, Studia Mathematica 279 (2024), 1-69

[4] Candela, González-Sánchez, and Szegedy, *Spectral algorithms in higher-order Fourier analysis*, https://arxiv.org/abs/2501.12287

[5] Candela, González-Sánchez, and Szegedy, *On measure-preserving $\mathbb{F}_p^\omega$-systems of order k* Ergodic Theory and Dynamical Systems, Journal dAnalyse Mathematique, 2024

[6] Candela, González-Sánchez, and Szegedy, *On the inverse theorem for Gowers norms in abelian groups of bounded torsion* Ergodic Theory and Dynamical Systems, https://arxiv.org/abs/2311.13899

[7] Candela, González-Sánchez, and Szegedy, *A refinement of Cauchy-Schwarz complexity*, Ergodic Theory and Dynamical Systems, European Journal of Combinatorics, Volume 106, December 2022, 103592

[8] Candela and Szegedy, *Regularity and inverse theorems for uniformity norms on compact abelian groups and nilmanifolds.* Journal für die reine und angewandte Mathematik (Crelles Journal), 2022(789):1–42, May 2022.

[9] Kunszenti, Lovász, Szegedy, *Subgraph densities in Markov spaces*, Advances in Mathematics Volume 437, February 2024, 109414

[10] Kunszenti, Lovász, Szegedy, *Random homomorphisms into the orthogonality graph*, Journal of Combinatorial Theory, Series B, Volume 167, July 2024, Pages 392-444

[11] Backhausz, Bordenave, Szegedy, *Typicality and entropy of processes on infinite trees*, Ann. Inst. H. Poincaré Probab. Statist. 58(4): 1959-1980 (November 2022). DOI: 10.1214/21-AIHP1233

[12] Terjék, González-Sánchez, *Optimal transport with f-divergence regularization and generalized Sinkhorn algorithm.* Proceedings of The 25th International Conference on Artificial Intelligence and Statistics, PMLR 151:5135-5165, 2022.

[13] Terjék, González-Sánchez, *A framework for overparametrized learning*, https://arxiv.org/pdf/2205.13507

[14] Szegedy, Czifra, Kőrösi-Szabó, *Dynamic memory based adaptive optimization*, https://arxiv.org/abs/2402.15262

[15] Szegedy, *A higher order gemneralization of group theory* https://arxiv.org/abs/2407.07815