RESEARCH STATEMENT IN FOUNDATIONS OF MACHINE LEARNING

MILTON LIN

Personal Experiences

My research in pure mathematics has centered on bridging discrete structures (number theory) with continuous spaces (topology) through *algebraic formalism*. Building on this, I aim to explore how **algebraic and categorical models** can reveal the **qualitative dynamics** of modern language models and neural circuits. This involves examining the extent to which **biological realism** can be mirrored in computational frameworks.

The overarching goal is to unify insights from *cognitive science*, *category theory*, and *computation* to understand the *limiting behaviors* of complex networks, both artificial and biological. Specifically, my research will offer new perspectives on *scaling laws* in large language models [Kap+20], [SMK23] and propose biologically grounded designs for neural networks. In Section 1, I introduce *associative memory networks*. My joint research studies the weight space decomposition of these models and evaluates their scaling behavior. In Section 2, I highlight how my background in algebra and geometry positions me to contribute to this interdisciplinary research.[1]

1. Associative Memory Networks

Associative memory networks, particularly Hopfield networks, were among the early computational models for memory search and retrieval [Kah20]. Recent developments have significantly advanced these models along two fronts: i) *Improved storage capacity*, progressing from polynomial [KH16], to exponential [Dem+17], and in other point of views, [HT14] ii) *Integration into modern deep learning architectures*, such as attention mechanisms [Ram+21], energy-based transformers [Hoo+23], and higher-order models like simplicial Hopfield networks [BF23]. Their relations with, and their potential to explain, modern transformer-based decoder models are under explored.

Research Goal: Scaling Properties of Associative Memory and Modern Models. The two key research areas are, joint with Chris Hillar (Redwood Research), Tenzin Chan (Algebraic) and Muhan Gao (Johns Hopkins University)

1. *Polytopal Decomposition of Weight Spaces in Toy Models*: We will extend the polytopal weight space decomposition, as present in literature on threshold linear networks, [CLM20], [CGM23], to higher order memory networks, such as *simplicial Hopfield networks* [BF23] or *dense associative memories*, [KH16]. This connects to recent approaches using spline theory to understand neural networks, [Bb18], [Bla+22]. We will study how the decomposition changes as the *network size increases*.

2. *Modern associative networks beyond memory capacity*: We will evaluate dense assocative memories, [KH16] beyond the theoretical memory capacity, see Equation (5) and (6) of [KH16]. While much effort has been focused on designing networks that extends the memory capacity, there is little work on studying such regimes. Our first empirical results show that storage capacity is not a hard constraint to task performance. Such insensitivity to memory capacity echoes trends seen in scaling laws of deep learning. Moving forward, by leveraging the interperable aspects of stored memory and energy land scape, we are exploring:

[1] I am indebted to my discussions with Prof. Daniel Khashabi and Prof. Leyla Isik.

(1) *Generalization and catastrophic forgetting:* The behavior of stored memory patterns appears highly sensitive to the nature of the task. How does task variability influence memory retrieval, and could this sensitivity offer insights into *catastrophic forgetting*? Understanding this phenomenon, especially in the context of continual learning, could bridge memory networks with advances in lifelong machine learning [Kem+17].

(2) *Correlated data and memory convergence* Experimental evidence shows that correlated datasets significantly alter convergence behavior to stored memory patterns. Can these observations be formalized theoretically? A deeper understanding of how data structure impacts memory retrieval could inform both theoretical bounds and practical applications.

(i) *Data augmentation:* We study the effect of such operations during training, on memory storage and retrieval efficiency. Our setup is similar to prior work, [AL24].

The end goal is to provide both empirical and theoretical comparison with modern networks; works along these lines include, [ND21], [Niu+24], and [CDB24].

2. Categorical Models and Homotopy Theory

Categorical approaches have gained momentum as a systematic framework for studying network structures [Gav+24]. This has been particularly successful in the field of *geometric deep learning* [Bro+21], where abstract mathematical structures help describe complex neural networks. We propose to explore *Hopfield networks* using a recent formalism by Manin et al. [MM24], which uses *summing functors* and *Gamma spaces* to model the allocation of resources in neural networks. These concepts will allow us to understand how the complexity of memory networks scales as network size increases. The formalism allows us to study a *homotopy type* - a mathematical construct at a deeper level than *homology*[2]. Homotopy captures invariants of network up to continuous deformations. Previous studies have shown that stimulus space can be reconstructed up to homotopy [Man15].

Research Goal: Homotopical Complexity Under Scaling. This research will investigate how the **homotopical complexity** of memory networks evolves as their size increases. Specifically, we will examine how *memory capacity* correlates with *homotopical invariants* like *Betti numbers* (which measure the number of independent cycles in a space) and *simplicial complexes* (which provide a higher-dimensional generalization of networks). Burns and Fukai have already done early work in this direction [BF23], but much remains to be explored.

3. Expected impact:

These research will highlight

- the limitations of synthetic memory networks, particularly in the case when the number of parameters exceed the storage capacities. This sheds light in regarding these networks as proxies for explaining biological networks, see also [KH21].
- the possibilities of creating hybrid models that respect biological constraints while maintaining the computational power of synthetic networks.
- lens through which to study the behavior of neural networks algebraically.

[2] which is commonly used in topological data analysis (TDA). For a short survey of topology and neural code, see [Cur16].

References

[AL24] Allen-Zhu, Zeyuan and Li, Yuanzhi. *Physics of Language Models: Part 3.1, Knowledge Storage and Extraction*. 2024. arXiv: 2309.14316 [cs.CL]. URL: https://arxiv.org/abs/2309.14316 (cit. on p. 2).

[Bb18] Balestriero, Randall and baraniuk, richard. "A Spline Theory of Deep Learning". In: *Proceedings of the 35th International Conference on Machine Learning*. Ed. by Jennifer Dy and Andreas Krause. Vol. 80. Proceedings of Machine Learning Research. PMLR, Oct. 2018, pp. 374–383. URL: https://proceedings.mlr.press/v80/balestriero18b.html (cit. on p. 1).

[BF23] Burns, Thomas F and Fukai, Tomoki. "Simplicial Hopfield networks". In: *The Eleventh International Conference on Learning Representations*. 2023. URL: https://openreview.net/forum?id=_QLsH8gatwx (cit. on pp. 1, 2).

[Bla+22] Black, Sid, Sharkey, Lee, Grinsztajn, Leo, Winsor, Eric, Braun, Dan, Merizian, Jacob, Parker, Kip, Guevara, Carlos Ramón, Millidge, Beren, Alfour, Gabriel, and Leahy, Connor. *Interpreting Neural Networks through the Polytope Lens*. 2022. arXiv: 2211.12312 [cs.LG]. URL: https://arxiv.org/abs/2211.12312 (cit. on p. 1).

[Bro+21] Bronstein, Michael M., Bruna, Joan, Cohen, Taco, and Veličković, Petar. *Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges*. 2021. arXiv: 2104.13478 [cs.LG]. URL: https://arxiv.org/abs/2104.13478 (cit. on p. 2).

[CDB24] Cabannes, Vivien, Dohmatob, Elvis, and Bietti, Alberto. *Scaling Laws for Associative Memories*. 2024. arXiv: 2310.02984 [stat.ML]. URL: https://arxiv.org/abs/2310.02984 (cit. on p. 2).

[CGM23] Curto, Carina, Geneson, Jesse, and Morrison, Katherine. *Stable fixed points of combinatorial threshold-linear networks*. 2023. arXiv: 1909.02947 [q-bio.NC]. URL: https://arxiv.org/abs/1909.02947 (cit. on p. 1).

[CLM20] Curto, Carina, Langdon, Christopher, and Morrison, Katherine. *Combinatorial Geometry of Threshold-Linear Networks*. 2020. arXiv: 2008.01032 [math.CO]. URL: https://arxiv.org/abs/2008.01032 (cit. on p. 1).

[Cur16] Curto, Carina. *What can topology tell us about the neural code?* 2016. arXiv: 1605.01905 [q-bio.NC]. URL: https://arxiv.org/abs/1605.01905 (cit. on p. 2).

[Dem+17] Demircigil, Mete, Heusel, Judith, Löwe, Matthias, Upgang, Sven, and Vermet, Franck. "On a Model of Associative Memory with Huge Storage Capacity". In: 168 (May 2017), pp. 288–299 (cit. on p. 1).

[Gav+24] Gavranović, Bruno, Lessard, Paul, Dudzik, Andrew, Glehn, Tamara von, Araújo, João G. M., and Veličković, Petar. *Position: Categorical Deep Learning is an Algebraic Theory of All Architectures*. 2024. arXiv: 2402.15332 [cs.LG]. URL: https://arxiv.org/abs/2402.15332 (cit. on p. 2).

[Hoo+23] Hoover, Benjamin, Liang, Yuchen, Pham, Bao, Panda, Rameswar, Strobelt, Hendrik, Chau, Duen Horng, Zaki, Mohammed J., and Krotov, Dmitry. *Energy Transformer*. 2023. arXiv: 2302.07253 [cs.LG]. URL: https://arxiv.org/abs/2302.07253 (cit. on p. 1).

[HT14] Hillar, Christopher J. and Tran, Ngoc Mai. "Robust Exponential Memory in Hopfield Networks". In: *Journal of Mathematical Neuroscience* 8 (2014). URL: https://api.semanticscholar.org/CorpusID:11295055 (cit. on p. 1).

[Kah20] Kahana, Michael J. "Computational Models of Memory Search." In: *Annual review of psychology* (2020). URL: https://api.semanticscholar.org/CorpusID:203624267 (cit. on p. 1).

[Kap+20] Kaplan, Jared, McCandlish, Sam, Henighan, Tom, Brown, Tom B., Chess, Benjamin, Child, Rewon, Gray, Scott, Radford, Alec, Wu, Jeffrey, and Amodei, Dario. *Scaling*

Laws for Neural Language Models. 2020. arXiv: 2001.08361 [cs.LG]. URL: https://arxiv.org/abs/2001.08361 (cit. on p. 1).

[Kem+17] Kemker, Ronald, Abitino, Angelina, McClure, Marc, and Kanan, Christopher. "Measuring Catastrophic Forgetting in Neural Networks". In: *ArXiv* abs/1708.02072 (2017). URL: https://api.semanticscholar.org/CorpusID:22910766 (cit. on p. 2).

[KH16] Krotov, Dmitry and Hopfield, John J. *Dense Associative Memory for Pattern Recognition*. 2016. arXiv: 1606.01164 [cs.NE]. URL: https://arxiv.org/abs/1606.01164 (cit. on p. 1).

[KH21] Krotov, Dmitry and Hopfield, John. *Large Associative Memory Problem in Neurobiology and Machine Learning*. 2021. arXiv: 2008.06996 [q-bio.NC]. URL: https://arxiv.org/abs/2008.06996 (cit. on p. 2).

[Man15] Manin, Yuri I. *Neural codes and homotopy types: mathematical models of place field recognition*. 2015. arXiv: 1501.00897 [math.HO]. URL: https://arxiv.org/abs/1501.00897 (cit. on p. 2).

[MM24] Manin, Yuri and Marcolli, Matilde. "Homotopy Theoretic and Categorical Models of Neural Information Networks". In: *Compositionality* Volume 6 (2024) (Sept. 2024). ISSN: 2631-4444. URL: http://dx.doi.org/10.46298/compositionality-6-4 (cit. on p. 2).

[ND21] Nichol, Alex and Dhariwal, Prafulla. *Improved Denoising Diffusion Probabilistic Models*. 2021. arXiv: 2102.09672 [cs.LG]. URL: https://arxiv.org/abs/2102.09672 (cit. on p. 2).

[Niu+24] Niu, Xueyan, Bai, Bo, Deng, Lei, and Han, Wei. *Beyond Scaling Laws: Understanding Transformer Performance with Associative Memory*. 2024. arXiv: 2405.08707 [cs.LG]. URL: https://arxiv.org/abs/2405.08707 (cit. on p. 2).

[Ram+21] Ramsauer, Hubert, Schäfl, Bernhard, Lehner, Johannes, Seidl, Philipp, Widrich, Michael, Gruber, Lukas, Holzleitner, Markus, Adler, Thomas, Kreil, David, Kopp, Michael K, Klambauer, Günter, Brandstetter, Johannes, and Hochreiter, Sepp. "Hopfield Networks is All You Need". In: *International Conference on Learning Representations*. 2021. URL: https://openreview.net/forum?id=tL89RnzIiCd (cit. on p. 1).

[SMK23] Schaeffer, Rylan, Miranda, Brando, and Koyejo, Sanmi. *Are Emergent Abilities of Large Language Models a Mirage?* 2023. arXiv: 2304.15004 [cs.AI]. URL: https://arxiv.org/abs/2304.15004 (cit. on p. 1).