RESEARCH STATEMENT

MILTON LIN

My research[1] in the geometric Langlands program has centered on bridging discrete structures (number theory) with continuous spaces (topology) through *algebraic formalism.* I discuss this in Section 1. Building on this experience, 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.

$$
\begin{tikzcd}
& \text{Categorical} \arrow[dl, bend right] \arrow[dr, bend left] & \\
\text{Cognitive} \arrow[rr, leftrightarrow, bend right] & & \text{Computational}
\end{tikzcd}
$$

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.

1. Past Research: p-adic Geometry and Langlands Program

My graduate work[2] has focused on extensions in the mixed characteristic setting, [ILZ24], this is joint work with Ashwin Iyengar (American Mathematical Society) and Konrad Zou (Bonn Univeristy). Our work applied the framework of Zhu's perfect geometry [Zhu17] to prove the Casselman-Shalika formula [NP01] for mixed characteristics. The Casselman-Shalika formula computes the "fourier coefficients" of automorphic forms and is fundamental to modern works of geometric Langlands, see [FR22]. Moving forward, I will continue this research in two directions:

(1) **Relative aspects of Langlands**, joint project with Yuta Takaya (University of Tokyo), we aim to explore relative aspects of the Langlands program on the Fargues-Fontaine curve, [FS24], recent conjectures of Ben-Zvi. Sakellaridis, and Venkatesh [BSV], particularly the relationship between period sheaves and L-sheaves as in [FW24].

(2) **Metaplectic aspects of Langlands**, joint with Toan Pham (Johns Hopkins Univeristy) I intend to give a geometric metaplectic Casselman–Shalika formula, building on the works of Gaitsgory and Lysenko [GL22], McNamara [McN16], and Brubaker et al. [Bru+24].

Expected Impact: The research contributes to

Date: November 1, 2024.

[1] My mathematical research statement is available at https://cwlin4916.github.io/Trees/Application/Postdoc/Research.pdf.

[2] Details of my mathematical research: https://cwlin4916.github.io/Trees/Application/Postdoc/Research.pdf

- the fundamental study of representation theory, which appears in quantum physics to modern day machine learning, [Bro+21].
- foundations of the geometric Langlands program, in particular, to various aspects of relative Langlands, [BSV], that describes deep relations between the arithmetic and geometric spaces.

2. Machine Learning: Geometry of Associative Memory Networks

In machine learning, I am particularly interested[3] in foundational theories of associative memory networks, going back to the work of Hopfield and to modern-day associative memories, [KH16]. These networks serve as a bridge between biological realism and computational efficiency.

(1) **Polytopal Decomposition of Memory Networks** in joint work with Chris Hillar (former Redwood Research, currently startup on Algebraic) and Tenzin Chan (Algebraic) we focus on the polytopal decomposition of the weight spaces of memory networks and its relation to network scaling. Similar works include, [ZNL18].

(2) **Study of Homotopy Type:** To further my study with Chris Hillar, we propose to explore *Hopfield networks* using a recent formalism by Manin and Marcolli [MM24], which uses *summing functors* and *Gamma spaces* to model the allocation of resources in neural networks. The formalism allows us to study a *homotopy type* - a mathematical construct at a deeper level than *homology*[4]. Homotopy captures invariants of the network up to continuous deformations.

(3) **Interpretability of Energy Transformers** My joint work with Muhan Gao (Johns Hopkins University) empirically studies modern energy-based memory network transformers for language modeling. This was previously studied in the context of vision task in [Hoo+24]. Due to the interpretable nature of the networks, through their memory module, we are currently studying the representations learnt by these models, exploring the *platonic representation hypothesis* [Huh+24] which posits that modern neural networks are converging to a shared statistical model of reality in their representation space.

Expected impact: This 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.

3. Timeline for Accomplishments

In the first year of my fellowship, I will finalize my joint work with Chris Hillar (Redwood Research) on the polytopal decomposition of memory networks and with Muhan Gao (Johns Hopkins University) on experiments involving memory networks in language modeling. I also aim to develop theoretical results supporting associative memory models and their scaling properties, striving for

[3]Details of my interest in memory networks https://cwlin4916.github.io/Trees/Application/Postdoc/ResearchA.pdf

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

both mathematical rigor and practical applications in AI. Concurrently, I hope to continue researching foundational mathematics in the geometric Langlands program, completing my own and joint research.

References

[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).

[Bru+24] Brubaker, Ben, Buciumas, Valentin, Bump, Daniel, and Gustafsson, Henrik P. A. *Metaplectic Iwahori Whittaker functions and supersymmetric lattice models*. 2024. arXiv: 2012.15778 [math.RT]. URL: https://arxiv.org/abs/2012.15778 (cit. on p. 1).

[BSV] Ben-Zvi, David, Sakellaridis, Yiannis, and Venkatesh, Akshay. "Relative Langlands duality". In: () (cit. on pp. 1, 2).

[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).

[FR22] Faergeman, Joakim and Raskin, Sam. *Non-vanishing of geometric Whittaker coefficients for reductive groups*. 2022. arXiv: 2207.02955 [math.RT]. URL: https://arxiv.org/abs/2207.02955 (cit. on p. 1).

[FS24] Fargues, Laurent and Scholze, Peter. "Geometrization of the local Langlands correspondence". In: *arXiv e-prints*, arXiv:2102.13459 (Feb. 2024), arXiv:2102.13459. arXiv: 2102.13459 [math.RT] (cit. on p. 1).

[FW24] Feng, Tony and Wang, Jonathan. *Geometric Langlands duality for periods*. 2024. arXiv: 2402.00180 [math.NT]. URL: https://arxiv.org/abs/2402.00180 (cit. on p. 1).

[GL22] Gaitsgory, D. and Lysenko, S. *Parameters and duality for the metaplectic geometric Langlands theory*. 2022. arXiv: 1608.00284 [math.AG]. URL: https://arxiv.org/abs/1608.00284 (cit. on p. 1).

[Hoo+24] Hoover, Benjamin, Strobelt, Hendrik, Krotov, Dmitry, Hoffman, Judy, Kira, Zsolt, and Chau, Duen Horng. *Memory in Plain Sight: Surveying the Uncanny Resemblances of Associative Memories and Diffusion Models*. 2024. arXiv: 2309.16750 [cs.LG]. URL: https://arxiv.org/abs/2309.16750 (cit. on p. 2).

[Huh+24] Huh, Minyoung, Cheung, Brian, Wang, Tongzhou, and Isola, Phillip. *The Platonic Representation Hypothesis*. 2024. arXiv: 2405.07987 [cs.LG]. URL: https://arxiv.org/abs/2405.07987 (cit. on p. 2).

[ILZ24] Iyengar, Ashwin, Lin, Milton, and Zou, Konrad. *Geometric Casselman-Shalika in mixed characteristic*. 2024. arXiv: 2408.07953 [math.AG]. URL: https://arxiv.org/abs/2408.07953 (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).

[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. 2).

[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).

[McN16] McNamara, Peter J. "The metaplectic Casselman-Shalika formula". In: *Trans. Amer. Math. Soc.* 368.4 (2016), pp. 2913–2937. ISSN: 0002-9947,1088-6850. URL: https://doi.org/10.1090/tran/6597 (cit. on p. 1).

[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).

[NP01] Ngô, B. C. and Polo, P. "Résolutions de Demazure affines et formule de Casselman-Shalika géométrique". In: *J. Algebraic Geom.* 10.3 (2001), pp. 515–547. ISSN: 1056-3911,1534-7486 (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).

[Zhu17] Zhu, Xinwen. "Affine Grassmannians and the geometric Satake in mixed characteristic". In: *Ann. of Math. (2)* 185.2 (2017), pp. 403–492. ISSN: 0003-486X,1939-8980. URL: https://doi.org/10.4007/annals.2017.185.2.2 (cit. on p. 1).

[ZNL18] Zhang, Liwen, Naitzat, Gregory, and Lim, Lek-Heng. *Tropical Geometry of Deep Neural Networks*. 2018. arXiv: 1805.07091 [cs.LG]. URL: https://arxiv.org/abs/1805.07091 (cit. on p. 2).