RESEARCH STATEMENT

MILTON LIN

My research revolves around foundational aspects of pure mathematics and machine learning. In mathematics, I concentrate on the geometric Langlands program, mainly its metaplectic and relative extensions. My recent interest in machine learning is exploring associative memory models and scaling properties through topological and algebraic methods. The first two pages of this document provide a summary. I have provided hyperlinks for details.

Research in Mathematics: p-adic Geometry and the Langlands Program

I briefly describe two main pure mathematics projects I aim to complete by the Summer 2025. These are two of my important recent works. [1] In the geometric Langlands program, my graduate work has focused on extensions in the mixed characteristic setting, where joint with Ashwin Iyengar (American Mathematical Society) and Konrad Zou (Bonn Univeristy). [ILZ24] we applied the framework of Zhu's perfect geometry [Zhu17] to prove the Casselman-Shalika formula in mixed characteristics [ILZ24]. 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) **Metaplectic aspects of Langlands**, see Section 3 for details, 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].

(2) **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].

Research in Machine Learning: Geometry of Associative Memory Networks

In machine learning, I am particularly interested 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. I hope to apply my background in algebra, category theory, and geometry to give insights into the nature of modern networks. A brief summary of my current projects are:

(1) **Polytopal Decomposition of Memory Networks** joit with Chris Hillar (Redwood Research) we focus on the polytopal decomposition of the weight spaces of memory networks and its relation to network scaling. Similar works include, [ZNL18]. In the future, we hope to explore these networks using the recent formalism by Manin and Marcolli [MM24]. This potentially allows - if possible - us to study the relationship between memory capacity and homotopical invariants.

Date: December 16, 2024.

[1]Details : https://cwlin4916.github.io/Trees/Application/Postdoc/Research.pdf

(2) **Dense Associative Memories beyond the storage capcity** joint with Muhan Gao (Johns Hopkins University) we study dense associative memories [KH16] and its variations [Hoo+23] for language modeling and classification tasks. We study the regime where the stored memories are beyond the theoretical capacity; see Equation (5) and (6) of [KH16]. This research will highlight the limitations of synthetic memory networks, especially in their use as proxies for explaining biological networks, see [KH21].

We refer to Section 2 for more details.

1. Future work and Career

In the future, I aim to transition my research focus toward the theoretical foundations of machine learning and biologically plausible models, leveraging my mathematical expertise. By the summer of 2025, I plan to complete my ongoing research with Chris Hillar on neural network parameter space decompositions and finalize my work in pure mathematics.

During the summer and fall of 2025, I will further expand my expertise by working on reinforcement learning and language models at the Center for Human-Compatible Artificial Intelligence, UC Berkeley, under Michael K. Cohen.

At the Flatiron Institute, I am particularly enthusiastic about engaging in collaborative research. I am drawn to the work of Professors Michael Eickenberg and Alberto Bietti from the Center for Computational Mathematics, as well as Dmitri Chklovskii, Alex Williams, David Lipshutz, and Siavash Golkar from the Center for Computational Neuroscience. These collaborations would enable me to explore intersections between computational neuroscience, machine learning, and advanced mathematical frameworks.

2. 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: Polytopal Weight Decomposition in Neural Networks. My current research investigates the decomposition of neural network parameter spaces using polytopal and tropical geometric methods, with a primary focus on associative memory networks. This work, conducted in collaboration with Chris Hillar (Redwood Research, Berkeley), builds upon foundational studies of hyperplane arrangements in neural networks [PMB14; Mon+14; Xio+20] and tropical geometry [CM19].

We analyze the complexity and dynamics of memory networks by studying the parameter space $\mathrm{Par} = \mathbb{R}^{n(n+1)}$, where $\Theta = (W, b) \in \mathrm{Par}$ consists of weights $W \in \mathbb{R}^{n\times n}$ and biases $b \in \mathbb{R}^n$. These parameters define the update function of an associative memory network:

$$\mathrm{HN}_\Theta(x) = H(Wx + b) : \{0,1\}^n \to \{0,1\}^n,$$

where H is the Heaviside function, defined by $H(r) = 1$ if $r > 0$ and $H(r) = 0$ otherwise. Repeated applications of this map yield higher iterations:

$$\mathrm{HN}_\Theta^{\circ k} = \underbrace{\mathrm{HN}_\Theta \circ \cdots \circ \mathrm{HN}_\Theta}_{k \text{ times}}.$$

While prior work has focused on the polyhedral decomposition of the domain space of HN_Θ for fixed Θ, our research emphasizes the decomposition of the parameter space Par itself. Specifically, for a fixed function $f \in \mathrm{Fct}(\{0,1\}^n, \{0,1\}^n)$, we define the parameter set:

$$\mathrm{Par}_f^k = \{\Theta \in \mathrm{Par} \mid \mathrm{HN}_\Theta^{\circ k} = f\} \subset \mathrm{Par}.$$

This forms a polyhedron in the parameter space, and the collection of such polyhedra induces a simplicial hyperplane arrangement [Zie94], denoted $\Sigma[\mathrm{Par}^k]$. The structure of $\Sigma[\mathrm{Par}^k]$ is central to our analysis, as it captures the functional expressivity of Hopfield networks after k iterations. The key research questions we explore are:

- How does the simplicial complex $\Sigma[\mathrm{Par}^k]$ evolve as k increases - and how do such results compare with infinite width theory [BP23]? This exploration sheds light on the iterative dynamics of recurrent neural networks.
- Can the combinatorial and topological properties of $\Sigma[\mathrm{Par}^k]$ provide insights into the convergence behavior of memory networks?
- What is the relationship between the parameter space decomposition and the polyhedral decomposition of the domain space induced by neural networks? Preliminary observations suggest a duality, as hinted in [Liu+23].

We aim to extend the decomposition framework to higher-order memory networks, such as *simplicial Hopfield networks* [BF23] and *dense associative memories* [KH16]. These extensions involve studying the domain space decomposition for these networks, with a focus on how the decomposition evolves as network size increases. This connects to recent work leveraging spline theory to analyze neural networks [BB18; Bla+22].

In conclusion, this research aims to establish a unifying framework for understanding recurrent neural networks through polyhedral and algebraic geometry. The insights gained could advance our understanding of network equivalence, capacity, and convergence, with applications in theoretical and applied machine learning.

Research Goal: Scaling Properties of Associative Memory and Modern Models. This joint work with Muhan Gao (Johns Hopkins University). 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 extend the memory capacity, there is little work on studying such regimes. see [Kal+24] for another perspective. 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) *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 relates memory capacity and retrieval could inform both theoretical bounds and practical applications.

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

Research Goal: Applying categorical perspectives 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]. In the future, we hope to explore scaling properties and memory capacity of associative memory networks with recent formalism of Manin and Marcolli [MM24], which uses *summing functors* and *Gamma spaces* to model the allocation of resources in neural networks. Burns and Fukai have done early work [BF23], but much remains to be explored.

3. Mixed Characteristic Geometry and the Casselman–Shalika formula

In this section, I provide a brief summary of my research. Feel free to omit this section. Let G be a connected reductive group over a nonarchimedian local field with residue characteristic $p \neq \ell$, and $\Lambda := \overline{\mathbb{Q}}_\ell$. In [FS24], Fargues and Scholze have formulated the geometric Langlands conjecture for the Fargues Fontaine curve, similar to the function field correspondence of Beilinson-Drinfeld [AG15], [Gai14]. One fundamental aspect of the program is to understand the *Whittaker Fourier coefficient functor*,

$$\mathrm{coeff} : D_{\mathrm{lis}}(\mathrm{Bun}_G, \Lambda) \to D(\Lambda)$$

and its various properties, see [FR22, Ch. 5] for definitions and theorems. In the classical setting, as described *op.cit.*, this corresponds to finding the Fourier coefficients of automorphic functions:

Example 3.1. Let $G = \mathrm{PGL}_2$ be the projective linear group over $\mathbb{Q}$. A modular function, f, has an adelic formulation, $\widetilde{f}$ on $G(\mathbb{A}_\mathbb{Q})$. Up to a *normalizing factor*, the Fourier coefficients coincide with the Whittaker coefficients (Jacquet period),

$$a_m \sim \int_{N(\mathbb{Q})\backslash N(\mathbb{A}_\mathbb{Q})} \widetilde{f}(n\alpha_m)\psi(-n)\, dn \qquad \text{for } m \geq 1$$

where $\alpha_m \in T(\mathbb{A}_\mathbb{Q}^{\mathrm{fin}})$ is m considered as a finite idèle and ψ is a standard character on $N(\mathbb{A}_\mathbb{Q}) = \mathbb{A}_\mathbb{Q}$, where N is the radical of standard Borel. For more details, see [Gel75, Ch.3].

The first fundamental result in this context is the *global Casselman-Shalika formula*, as proven in [FGV01], which joint Ashwin Iyengar (American Mathematical Society) and Konrad Zou (Bonn University) [ILZ24], we replicated a variation: the *geometric Casselman–Shalika formula* over the Witt vector affine Grassmannian Gr_G.

Theorem 3.2. *[ILZ24] The geometric Casselman–Shalika formula holds over the Witt vector Grassmanian.*

At the same time, we have established basic properties of averaging functors (cf. [FR22, Ch. 7], [BBM21]), yielding Iwahori–Satake equivalence, [Bez+19] without using nearby cycles.

Theorem 3.3 (I.-Lin-Z., in progress)**.** *The Iwahori-Whittaker category is equivalent to the spherical Hecke category in mixed characteristics.* [2]

Research Goal A. Geometrization of Metaplectic/Quasi-split Casselman–Shalika. *We propose two explorations of the Casselman–Shalika formula over equal-characteristic local fields. First, a geometric metaplectic Casselman–Shalika formula, building on the works of Gaitsgory and Lysenko [GL22], McNamara [McN16], and Brubaker et al. [Bru+20; Bru+24]. Second, a geometric Casselman–Shalika formula for quasi-split groups, following [GK20].*

[2] In the set up of [Bez+19], the equivalence was applied to modular representation theory, as the Iwahori-Whittaker category has a much simpler categorical structure.

References

[AG15] Arinkin, D. and Gaitsgory, D. “Singular support of coherent sheaves and the geometric Langlands conjecture”. In: *Selecta Math. (N.S.)* 21.1 (2015), pp. 1–199. ISSN: 1022-1824,1420-9020. URL: https://doi.org/10.1007/s00029-014-0167-5 (cit. on p. 4).

[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. 3).

[BBM21] Bezrukavnikov, Roman, Braverman, Alexander, and Mirkovic, Ivan. *Some results about the geometric Whittaker model.* 2021. arXiv: math/0210250 [math.AG] (cit. on p. 4).

[Bez+19] Bezrukavnikov, Roman, Gaitsgory, Dennis, Mirković, Ivan, Riche, Simon, and Rider, Laura. “An Iwahori-Whittaker model for the Satake category”. In: *J. Éc. polytech. Math.* 6 (2019), pp. 707–735. ISSN: 2429-7100,2270-518X. URL: https://doi.org/10.5802/jep.104 (cit. on p. 4).

[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. 2–4).

[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. 3).

[BP23] Bordelon, Blake and Pehlevan, Cengiz. *Dynamics of Finite Width Kernel and Prediction Fluctuations in Mean Field Neural Networks.* 2023. arXiv: 2304.03408 [stat.ML]. URL: https://arxiv.org/abs/2304.03408 (cit. on p. 3).

[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. 4).

[Bru+20] Brubaker, Ben, Buciumas, Valentin, Bump, Daniel, and Gustafsson, Henrik P. A. *Vertex operators, solvable lattice models and metaplectic Whittaker functions.* 2020. arXiv: 1806.07776 [math.RT]. URL: https://arxiv.org/abs/1806.07776 (cit. on p. 4).

[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 pp. 1, 4).

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

[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. 4).

[CM19] Charisopoulos, Vasileios and Maragos, Petros. *A Tropical Approach to Neural Networks with Piecewise Linear Activations.* 2019. arXiv: 1805.08749 [stat.ML]. URL: https://arxiv.org/abs/1805.08749 (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. 2).

[FGV01] Frenkel, E., Gaitsgory, D., and Vilonen, K. “Whittaker patterns in the geometry of moduli spaces of bundles on curves”. In: *Ann. of Math. (2)* 153.3 (2001), pp. 699–748. ISSN: 0003-486X,1939-8980. URL: https://doi.org/10.2307/2661366 (cit. on p. 4).

[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 pp. 1, 4).

[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 pp. 1, 4).

[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).

[Gai14] Gaitsgory, Dennis. *Outline of the proof of the geometric Langlands conjecture for GL(2)*. 2014. arXiv: 1302.2506 [math.AG]. URL: https://arxiv.org/abs/1302.2506 (cit. on p. 4).

[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. 4).

[Gel75] Gelbart, Stephen S. *Automorphic forms on adèle groups*. Vol. No. 83. Annals of Mathematics Studies. Princeton University Press, Princeton, NJ; University of Tokyo Press, Tokyo, 1975, pp. x+267 (cit. on p. 4).

[GK20] Gurevich, Nadya and Karasiewicz, Edmund. *The Twisted Satake Transform and the Casselman-Shalika Formula for Quasi-Split Groups*. 2020. arXiv: 2012.09893 [math.RT]. URL: https://arxiv.org/abs/2012.09893 (cit. on p. 4).

[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 pp. 1, 4).

[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. 2).

[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. 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 pp. 1, 4).

[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. 2).

[Kal+24] Kalaj, Silvio, Lauditi, Clarissa, Perugini, Gabriele, Lucibello, Carlo, Malatesta, Enrico M., and Negri, Matteo. *Random Features Hopfield Networks generalize retrieval to previously unseen examples*. 2024. arXiv: 2407.05658 [cond-mat.dis-nn]. URL: https://arxiv.org/abs/2407.05658 (cit. on p. 3).

[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. 3).

[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 pp. 1–3).

[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).

[Liu+23] Liu, Yajing, Cole, Christina M, Peterson, Chris, and Kirby, Michael. *ReLU Neural Networks, Polyhedral Decompositions, and Persistent Homology*. 2023. arXiv: 2306.17418 [math.AT]. URL: https://arxiv.org/abs/2306.17418 (cit. on p. 3).

[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 pp. 1, 4).

[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 pp. 1, 4).

[Mon+14] Montúfar, Guido, Pascanu, Razvan, Cho, Kyunghyun, and Bengio, Yoshua. *On the Number of Linear Regions of Deep Neural Networks*. 2014. arXiv: 1402.1869 [stat.ML]. URL: https://arxiv.org/abs/1402.1869 (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. 4).

[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. 4).

[PMB14] Pascanu, Razvan, Montufar, Guido, and Bengio, Yoshua. *On the number of response regions of deep feed forward networks with piece-wise linear activations*. 2014. arXiv: 1312.6098 [cs.LG]. URL: https://arxiv.org/abs/1312.6098 (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. 2).

[Xio+20] Xiong, H., Huang, L., Yu, M., Liu, L., Zhu, F., and Shao, L. *On the Number of Linear Regions of Convolutional Neural Networks*. 2020. arXiv: 2006.00978 [cs.LG]. URL: https://arxiv.org/abs/2006.00978 (cit. on p. 2).

[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).

[Zie94] Ziegler, Günter M. "Lectures on Polytopes". In: 1994. URL: https://api.semanticscholar.org/CorpusID:117286447 (cit. on p. 3).

[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. 1).