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, particularly its metaplectic and relative extensions. My recent interests in machine learning explore associative memory models and scaling properties, through topological and algebraic methods. The first three pages of this document provide a summary, followed by details.

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

I split my current and future projects into two categories: **core projects**, where I am primarily focused on advancing mixed characteristic and metaplectic aspects of the Langlands program, and additional **ongoing work in related areas**, including categorical deformations.

Core proejcts. 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 1 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**, see Section 2 for details, 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].

Related works. In addition to my primary projects I am equally committed to the broader foundation of representation theory:

(1) **Motivic aspects of Langlands** , see Section 3 for details, Building on [RS20], I aim to define a Whittaker category in motivic setting, obtaining a similar equivalence at [FGV01]. To begin, we prove the same statement in [NP01] in the category of mixed Tate motives. The difficulty is that the original argument requires characteristic 0 coefficients, whilst in positive characteristic, there are problems of extensions.

(2) **Stacky Approaches to Periods**, see Section 4 for details, I use recent advancements in formalism of stacks to study periods. I have proven that the unipotent fundamental group associated to a pointed scheme can be recovered via a *stacky approach*, see [GR14], [Toë06], [MR23]: for a given scheme X, there exists a natural stack, $\mathbf{U}(X_{\mathrm{dR}})$ whose fundamental group coincides with the unipotent de Rham fundamental group as studied in [Bro14]. By similar techniques of [Bha23], we should recover Haine's theorem: the pro-unipotent

completion of de Rham fundamental group admits a mixed Hodge structure. [1] We hope that such work can spark new techniques and new phenomena, such as those used in p-adic integration theory, [Vol01].

(3) **Categorical deformations of representation category**, see Section 5 for details, this builds upon my current research on the Whittaker category, from the point of view of deformation theory. We will first document a careful proof of Lurie's theorem, [Lur10, Thm 10.10], which describe *formal deformation of categories*, as gerbes see [Lur10, Ch.8-10] for definitions. Then, we will explore deformations of representation of Lusztig's small quantum group, using recent advances in quantum geometric Langlands. Ultimately, this research aims to contribute to the foundations of categorical deformations.

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 aim 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) **Combinatorial Geometry of Parameter Space** joint with Chris Hillar (Redwood Research). We study the polytopal decomposition of the weight spaces of memory networks and interpret parameter estimation and network scaling under this perspective. Similar works include, [Mon+14], [ZNL18]. In the future, we hope to explore these networks using the recent formalism by Manin and Marcolli [MM24].

(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 6 for more details of the above two projects. Lastly, my focus of future endeavors is transitioning to reinforcement learning algorithms for mathematical research.

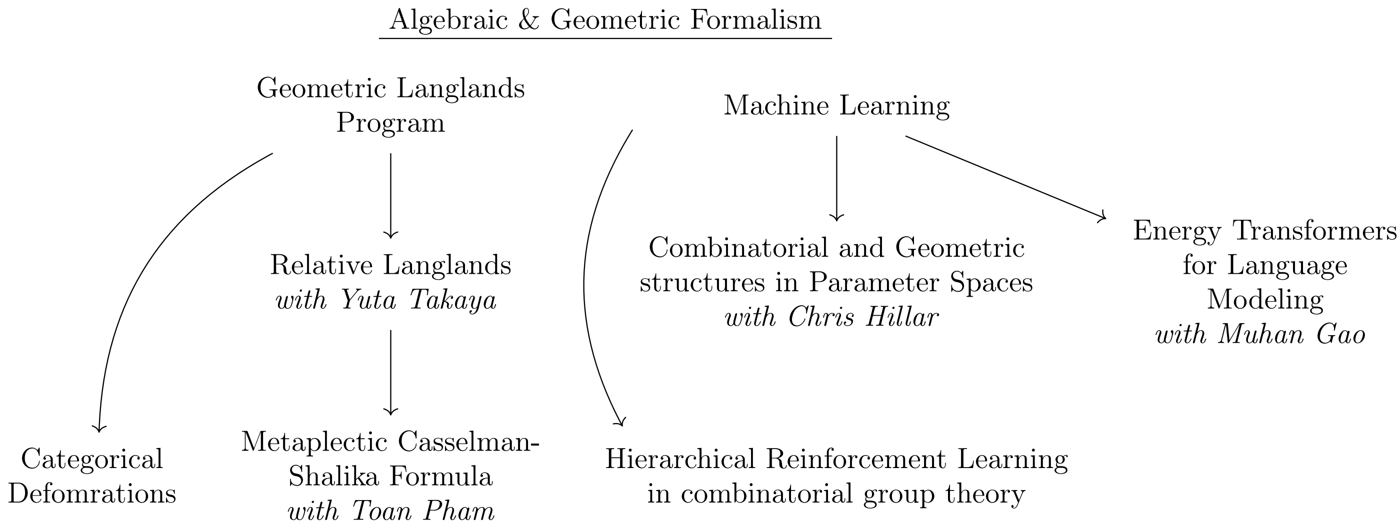
(3) **Reinforcement learning in mathematical research:** Building on a recent study of applications of reinforcement learning to mathematics - a research problem in combinatorial group theory, [She+24] - I hope to continue further and develop algorithms that expand the action space dynamically. As an initial step, I am implementing a variant of the options framework combined with hindsight experience replay [Lev+19] in the controlled experimental setting of *op. cit.*. The options framework of Markov decision processes stands as a first step in temporal planning, [SB18].

Suitability for the ICERM program and plan. I believe my background in geometric representation theory, number theory, and algebraic combinatorics aligns naturally with the program's focus on uncovering richer algebraic and geometric structures; my ongoing exploration of machine learning applications, developed just this last year, also complements the program's aim to integrate computational perspectives into these areas.

By 2025 fall, I will complete my collaboration with Yuta Takaya on relative Langlands for the Fargues–Fontaine curve in mathematics, see Section 1; study of geometric structures in parameter spaces of memory networks with Chris Hillar, see Section 6; and experiments with Muhan Gao on

[1] In Brown's approach, he reduced his study of motivic periods to mixed Hodge periods [Bro17, p. 3].

experiments involving memory networks in language modeling. Over the spring, I plan to extend my research on metaplectic aspects of Langlands and explore reinforcement learning algorithms for mathematics. To visually summarize my research timeline and collaborations, I present the following diagram: which includes the names of my co-authors.



The second row consists of projects I wish to complete by fall 2025 to spring 2026, and the third row is those that I wish to complete by summer 2026.

1. Mixed Characteristic Geometry and the Casselman–Shalika formula

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 1.1. 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),

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

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, analogous to the equi-characteristic geometrization carried out in [NP01].

Theorem 1.2. *[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 1.3 (I.-Lin-Z., in progress)**.** *The Iwahori-Whittaker category is equivalent to the spherical Hecke category in mixed characteristics.* [2]

As of writing, D. Hansen, L. Hamman, and L. Mann have announced significant advancements to the global version of the formula. Building on all these progresses, I have developed a strong interest in exploring the geometrization of the Casselman–Shalika formula for covering groups [GGW18].

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].*

Lastly, we would like to highlight that in our recent work [ILZ24], we have studied the smoothness of perfect Schubert subvarieties, $\mathrm{Gr}_{G,\leq\mu}$, for minuscule and quasi-minuscule μ. In the equal characteristic case over $\mathbb{C}((t))$, this was first examined by Evens and Mirkovic and later extended by Haines [HR20], with applications to classifying Shimura varieties with good or semi-stable reductions.

Research Goal B. Geometry of general perfect Schubert variety *Prove the results of Pappas and Zou, [PZ24] in perfect geometry. Associated to an absolutely special vertex in the Bruhats-Tits building of G, we have an associated group scheme $\mathcal{G}$ over $\mathcal{O}$.*

Conjecture 1.4. The smooth locus of $\mathrm{Gr}_{\mathcal{G},\leq\mathfrak{s}}$ is $\mathrm{Gr}_{\mathcal{G},\mathfrak{s}}$ in perfect geometry, in the sense of [Zhu17].

2. Relative langlands on the Fargues Fontaine curve

Joint with Yuta Takaya (Tokyo University), we explicitly compare the period sheaves on $\mathcal{A}$-side (automorphic) and L-sheaves on $\mathcal{B}$-side (Galois) under the relative Langlands conjectures of Ben-Zvi-Sakellaridis-Venkatesh. [BSV]. In *op. cit.*, one interprets Langlands correspondence as a form of arithmetic 4d quantum field theory, which has genesis in the Kapustin-Witten interpretation, [KW07], and the Knots and Primes promoted by B. Mazur, M. Kim, [Kim16]. The Fargues-Fontaine curve should be a *global object* of dimension 2 under the dictionary.

The theory suggests a duality extending that Langlands dual group: (G, X) with $(\hat{G}, \hat{X})$ of hyperspherical varieties. Let $\Lambda = \bar{\mathbb{Q}}_l$. We considered the *Iwasawa-Tate case*: $G = \mathbb{G}_{m,F}$ and $X = \mathbb{A}_F$ with dual pair $\hat{G} = \mathbb{G}_{m,\Lambda}$ and $\hat{X} = \mathbb{A}_\Lambda$. We constructed two maps

$$\pi : \mathrm{Bun}_G^X \to \mathrm{Bun}_G, \quad \hat{\pi} : \mathrm{LS}_{\hat{G}}^{\hat{X}} \to \mathrm{LS}_{\hat{G}}$$

yielding the *period sheaf*, $\mathcal{P}_X := \pi_!\Lambda$, and *L-sheaf*, $\mathcal{L}_{\hat{X}} := \hat{\pi}_*\omega_{\mathrm{Loc}_{\hat{G}}^{\hat{X}}}$. On the automorphic side, Bun_G has a Hardar-Narasimhan straification by locally closed substacks Bun_G^b indexed by the Kottwitz set $B(G) \simeq \mathbb{Z}$. Interesting phenomena occurs for $n \in \mathbb{Z}_{\geq 0} \subset \mathbb{Z} \simeq B(\mathbb{G}_m)$, and the study of period sheaves reduces to the study of the Abel-Jacobi map previously studied by Fargues [Far20] and Hansen [Han21]. On the other hand, we expect that one can study the spectral side using derived Fourier vector bundles, [FW24, Ch. 6] recently developed by Anschütz and Le Bras [AL21].

[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.

Conjecture 2.1. [Lin-T., in progress] Under the geometric local Langlands correspondence, $\mathbb{L}_G$, (appropriately normalized) $\mathcal{P}_X$ is sent to[3] $\mathcal{L}_{\hat{X}}$.

Research Goal C. Relative Langlands on the Fargues Fontaine curve. *Complete Conjecture 2.1 as a first step and then the Hecke case, which classically corresponds to Hecke's integral representation of standard L-function for GL_2. Lastly, one can ask whether on the $\mathcal{B}$-side, the same constructions of [BSV, Ch. 11] works for the p-adic (Emerton-Gee) L-parameter stacks, which potentially give new interpretations to p-adic L-functions.*

3. Motivic Phenomena

In this section S denotes a base scheme satisfying the Beilinson-Soulé vanishing conjecture, such as $\operatorname{Spec}\mathbb{Z}$, finite field. R denotes a regular coherent coefficient ring. I discuss one side project I have done, and one area of research, D, which extends my joint work [ILZ24].

In Grothendieck's quest towards the Weil conjectures, he reduced to problems of motives. Beilinson proposed an extension of the notion of pure motives to mixed motives, which was realized through the works of Voevodsky, Levine for a field, and extended by Ayoub and Cisikinki-Déglise, [CD19]. For a finite type S-scheme $X \in \mathrm{Sch}_S^{\mathrm{ft}}$, we can construct the derived categories of motives over X, $\mathrm{DM}(X, R)$. If X is smooth then its Ext-groups,

$$\pi_{-m} \operatorname{Map}_{\mathrm{DM}(X,R)}(1_X, 1_X(n)) \simeq \mathrm{CH}^n(X, 2n-m)_R$$

are Bloch's higher Chow groups. Motivic categories and the Chow groups are difficult to explicitly work with due to the lack of motivic t-structures. One method is due to the work of Beauville [Bea83], using Fourier transform. In my joint work, [Has+24] we extended the work of Beauville,

Theorem 3.1 (Lin et al.)**.** *Let $X \to Y$ be an abelian scheme, whose base Y is smooth and quasi projective over a field. There is an explicit N, such that one obtains a Beauville decomposition*

$$CH^i(X)_{\mathbb{Z}[1/N]} \simeq \bigoplus_s CH^i_{(s)}(X)_{\mathbb{Z}[1/N]}$$

where $CH^i_{(s)}(X)_R := \left\{x \in CH^i(X)_R \,:\, [n]_X^* x = n^{2i-s} x \quad \forall n \in \mathbb{Z}\right\}$ *and* $[n]$ *is the multiplication of an abelian scheme.*

This extends to a $\mathfrak{sl}_2$ action, which we discussed in *op.cit*; and if S were an algebraically closed field, this implies various structural results. The key ingredient was using G. Pappas' version of integral Grothendieck–Riemann–Roch, [Pap07].

Returning to motivic t-structures, it was shown by Levine that t-structures exist on a nice subcategory $\mathrm{DTM}(X) \hookrightarrow \mathrm{DM}(X)$ of *mixed Tate movies* for nice schemes X. This was extended to schemes with cellular Whitney-Tate stratification by Soergel and Wendt, and to prestacks in [RS20].

Research Goal D. Motivic Whittaker categories. *Define a Whittaker category in motivic setting, obtaining a similar equivalence at [FGV01]. To begin, we can prove the same statement in [NP01] in the category of mixed Tate motives, in which all the relevant objects are well defined. The difficulty, however, is that the original argument requires characteristic* 0 *coefficients, whilst in positive characteristic, there are problems of extensions.*

[3] One has to take into account shearing, twisting and tensoring. There is also an additional $\mathbb{G}_m := \mathbb{G}_{gr}$ action on X which we do not discuss.

4. Stacky approaches and periods

Various cohomology theories – crystalline cohomology, syntomic cohomology, and Dolbeaut cohomology – admit a factorization to the category of stacks over some affine scheme $\operatorname{Spec} R$,

$$\mathrm{Sch}_{\mathbb{Z}}^{\mathrm{sep,ft}} \to \mathrm{Stk}_R \to D(R)$$
$$X \mapsto X^? \mapsto \Gamma(X^?, \mathcal{O}_{X^?})$$

For instance, the *de Rham stack* X^{dR} over $\mathbb{Q}$, has points given by $X^{\mathrm{dR}}(A) := X(A_{\mathrm{red}})$ for any $\mathbb{Q}$-algebra A (cf. [GR14]). This is often referred to as a *stacky approach* [Dri22] or *transmutation* [Bha23], which allows one to use six functor formalism and geometric techniques. I explored this concept in basic constructions of F. Brown's work on *motivic periods*, [Bro14]. If X were a smooth variety over $\mathbb{Q}$, the matrix coefficient from Grothendieck's comparison theorem

$$H^*_{\mathrm{dR}}(X) \otimes_{\mathbb{Q}} \mathbb{C} \xrightarrow{\simeq} H^*_{\mathrm{Betti}}(X) \otimes_{\mathbb{Q}} \mathbb{C}$$

with respect to the $\mathbb{Q}$ structure of de-Rham and Betti cohomology (of $X(\mathbb{C})$) are *periods associated to* X. These periods along with their enhancements through Hodge structures, has a natural action of "Galois group"[4] which should govern the arithmetic structure of periods.

Let $X := \mathbb{P}^1 \setminus \{0, 1, \infty\}$ be the projective space minus three points over $\mathbb{Q}$. The period associated is studied through Chen's de Rham comparison theorem, [Che76], which relates to iterated integrals [Che73] and multiple zeta values [Bro14]. To give a "stacky" perspective of this period one uses unipotent types of stacks, which is defined in Toën, [Toë06], and recently in [MR23]. This is an endofunctor $\mathbf{U}$ on stacks, sending a stack X to its unipotent homotopy type. My first result is:

Theorem 4.1. *(Lin) Unipotent de Rham fundamental group,* $\pi_1^{u,dR}(X, x)$ *is isomorphic to* $\pi_1(\mathbf{U}(X^{dR}))$.

Research Goal E. A stacky approach to motivic periods. *This research aims to use geometric techniques in the study of periods. As a proof of concept, we will recover de Rham comparison theorem through the Riemann-Hilbert correspondence between the analytic Betti stack and the de Rham stack.* $\mathbb{P}^1 \setminus \{0, 1, \infty\}$*, is not proper, which requires us to incorporate log structures. I expect to prove:*

Conjecture 4.2. There exists X^{Betti} such that the unipotent Betti group $\pi_1^{u,\mathrm{Betti}}(X(\mathbb{C}), x)$ is isomorphic to $\pi_1(\mathbf{U}(X^{\mathrm{Betti}}))$. A logarithmic Riemann-Hilbert comparison should induce Chen's comparison theorem [Hai01, Thm 3.1]

$$\pi_1^{u,\mathrm{dR}}(X, x) \simeq \pi_1^{u,\mathrm{Betti}}(X(\mathbb{C}), x) \otimes \mathbb{C}$$

By similar techniques of [Bha23], we should recover Haine's theorem: the pro-unipotent completion of de Rham fundamental group admits a mixed Hodge structure. [5] *We hope that such work can spark new techniques and new phenomena, such as those used in p-adic integration theory, [Vol01].*

5. Deformation theory and the sphere spectrum $\mathbb{S}$

Let $\mathcal{S}$ denote the ∞-category of ∞-groupoids/anima. The *stabilization* of $\mathcal{S}$ is Sp, the ∞-category of spectra. This is the natural category to study cohomological invariants. Within Sp, lies the universal cohomology theory, $\mathbb{S}$, the *sphere spectrum* By Chevalley's works, connected reductive groups over $\mathbb{C}$ have a canonical split $\mathbb{Z}$-form $G_{\mathbb{Z}}$, see [Con15]. One can analogously ask: *is there a* $\mathbb{S}$*-form for algebraic groups*? A first approximation is the existence of an algebraic category $\mathrm{Rep}_{\mathbb{S}}(G_{\mathbb{S}})$, which

[4]For instance, in the approach of Deligne, he defined a *systems of realizations* [Del89]

[5]This is important in Brown's approach, where he reduced his study of motivic periods to mixed Hodge periods [Bro17, p. 3].

deforms to $\mathrm{Rep}_{\mathbb{Z}}(G_{\mathbb{Z}})$. To study, we begin with *formal deformation of categories*, which we briefly recall.

Let $\mathcal{C}$ be a symmetric monoidal category. There is a natural hierarchy of commutativity fitting in the diagram

$$\cdots \to \mathbb{E}_n(\mathcal{C}) \to \mathbb{E}_{n-1}(\mathcal{C}) \to \cdots \to \mathbb{E}_1(\mathcal{C})$$

where $\mathrm{CAlg}(\mathcal{C}) := \mathbb{E}_\infty(\mathcal{C}) := \varprojlim \mathbb{E}_n(\mathcal{C})$ of symmetric algebra objects can be identified with the limit. One can formalize the notion of $\mathbb{E}_n$-algebra objects via *disk operads*, or *little cubes operads*.

Example 5.1. Let $\mathcal{C} = (\mathcal{S}, \times)$. Let $Y \in \mathcal{S}_*$ is a pointed ∞-groupoid, its *k-fold based loop spaces*, $\Omega^k_* Y$ is a classical example of $\mathbb{E}_k$ algebra object in $(\mathcal{S}, \times)$.

Let $R \in \mathbb{E}_n(\mathrm{Sp})$ be an $\mathbb{E}_n$ ring, and consider LMod_R, the derived category of left R-modules, as an $\mathbb{E}_1$ object in $\mathrm{Pr}^{\mathrm{st}}$, the category of presentable stable categories. This defines $\mathrm{RMod}_{\mathrm{LMod}_R}(\mathrm{Pr}^{\mathrm{st}})$, the category of presentable stable (right) R-linear categories, [Lur18, Appendix D]. Set $\mathrm{Pr}^{\mathrm{st,cg}}_R$, as the full subcategory spanned by those whose underlying category is compactly generated. [6] For G a connected reductive group over a field k, $D^b(\mathrm{Rep}^{\mathrm{fd}}_k(G))$, the bounded derived category of finite dimensional algebraic representations with k coefficients lies in $\mathrm{Pr}^{\mathrm{st,cg}}_R$.

From now on, $k = \mathbb{C}$. Let $\mathrm{Art}^{(n)}_k$ denote the category of $\mathbb{E}_n$ *artinian* ring spectrum, for $n \geq 0$ over k. We refer to [Lur18, Ch. 15] for definition. This extends the classical definition of Artinian local ring, in particular, $R \in \mathrm{Art}^{(n)}_k$ admits an augmentation map $\epsilon : R \to k$. One defines the $\mathbb{E}_{n+2}$-formal moduli problem,

$$\mathrm{Cat}\mathcal{D}\mathrm{ef}^{(n)}(\mathcal{C}) : \mathrm{Art}^{(n+2)}_k \to \hat{\mathcal{S}}$$

$$R \mapsto |\, \{\mathcal{C}\} \times_{\mathrm{Pr}^{\mathrm{st,cg}}_R} \mathrm{Pr}^{\mathrm{st,cg}}_k |$$

where $|\quad|$ is the underlying Kan complex of the ∞-category. An object consists of: a $\mathcal{C}_B$ right stable R-linear category, and an equivalence $\mathcal{C}_B \otimes_{\mathrm{LMod}_B} \mathrm{LMod}_k \simeq \mathcal{C}$. Our $\mathbb{E}_4$-moduli problem is when $n = 2$ and $\mathcal{C} = D^b(\mathrm{Rep}^{\mathrm{fd}}_k(G))$. The geometric Casselman–Shalika [FGV01], which is the $\mathbb{E}_2$-algebra equivalence of $\mathcal{C}$ with the *Whittaker sheaves* on the *affine Grassmanian* $\mathrm{Gr}_{\hat{G}}$, describes this moduli problem. Consider moduli of functor of *$\mathbb{G}_m$-gerbes over X*

$$\mathrm{Ge}_{\mathbb{G}_m}(X) : R \mapsto \mathrm{Map}_{\mathbb{E}_2(\mathcal{S})}(X, B^2 R^\times) \quad R \in \mathrm{Art}^{(4)}_k$$

where $R^\times \subset \Omega^\infty R$ are the invertible elements of the underlying space of R [7] and B^2 is the second deloop. It was stated without proof in [Lur10]

Theorem 5.2 (Lurie)**.** *There is an equivalence of formal $\mathbb{E}_4$-moduli problems*

$$\widehat{Ge_{\mathbb{G}_m}}(\mathrm{Gr}_{\hat{G}}) \xrightarrow{\simeq} Cat\mathcal{D}ef^{(2)}(\mathrm{Rep}^{fd}_k G)$$

where $\widehat{-}$ is the formal completion of the moduli functor at a base point.

Research Goal F. Categorical deformations of the representation category *We will first document carefully Lurie's theorem, Theorem 5.2. Then, we will explore deformations of representation of Lusztig's small quantum group, as suggested in op.cit. Remark 10.12, using recent advances in quantum geometric Langlands. Ultimately, this research aims to contribute to the foundations of categorical deformations.*

---

[6] The compact generation is only a smallness condition for our version.

[7] is the union of the connected components of invertibles in the $\pi_0 R$ of the 0th space of R and is equivalent to the nth loop space of some space, $R^\times \simeq \Omega^n Z$, $n \geq 4$, hence the deloop $B^2 R^\times \in \mathbb{E}_2(\mathcal{S})$

6. Associative Memory Networks and Geometry of Parameter Spaces

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.

My first research goal is to study the geometry of parameter spaces of various networks. However, a particular emphasis has been placed on memory networks as they serve as recurrent networks compared to other feed-forward networks. The second research goal is to study more modern versions of memory networks - variations of dense associative memory networks.

Research Goal: Geometric structures in parameter space. Given a model architecture $\mathcal{A}$, such as transformers, CNN, multilayer perceptron, designed to interpolate a task $\mathcal{T}$, we investigate the encoded information within the parameter space $\mathrm{Par}_{\mathcal{A}}$. We define the mapping:

$$\mathcal{A}_{(-)} : \mathrm{Par}_{\mathcal{A}} \to \mathcal{T} \quad \Theta \mapsto \mathcal{A}_{\Theta}$$

which assigns each parameter Θ, $\mathcal{A}_{\Theta} \in \mathcal{T}$, an object of architecture $\mathcal{A}$, designed for a task, $\mathcal{T}$.

Example 6.1. Let $\mathcal{A} := \mathbb{F}[n, \sigma]$ be the class of L-layer feedforward neural network with hyperparameters: *width* $n = (n_i)_{i=1}^{L+1}$ and collection of *activation functions* σ $(\sigma_i : \mathbb{R}^{n_i} \to \mathbb{R}^{n_{i+1}})$.

For a set of a parameter

$$\Theta := \{A_i, b_i\}_{i=1}^{L} \in \mathbb{R}^{\sum_{i=1}^{L} n_i(n_i+1)}$$

we can associate a function

$$\mathcal{A}_{\Theta} := f_L \circ \cdots \circ f_1 : \mathbb{R}^{n_1} \to \mathbb{R}^{n_{L+1}}$$

For each $i = 1, \ldots, L$, and $f_i : \mathbb{R}^{n_i} \to \mathbb{R}^{n_{i+1}}$ is a linear function of the form

$$\sigma_i \left(A_i x_i + b_i\right),$$

This induces our desired map

$$\mathcal{A}_{(-)} : \mathrm{Par}_{\mathcal{A}} \to \mathrm{Fct}(\mathbb{R}^{n_1}, \mathbb{R}^{n_{L+1}})$$
$$\Theta \mapsto \mathcal{A}_{\Theta}$$

has, when $L \geq 2$, dense image within the subspace of continuous functions, [Cyb89].

The pair $(\mathcal{A}(-), \mathcal{T})$, partitions the collection of parameters according to those which induces the same object $f \in \mathcal{T}$. We obtain a collection of subsets of parameter space

$$\Sigma_{\mathcal{T}}[\mathrm{Par}_{\mathcal{A}}] = \{\mathrm{Par}_f\}_{f \in I(\mathcal{D})} \quad \mathrm{Par}_f = \{\Theta \in \mathrm{Par}_{\mathcal{A}} : \mathcal{A}_{\Theta} = f\}$$

by considering regions inducing the same object under the mapping $\mathcal{A}_{(-)}$. Often, $\Sigma_{\mathcal{T}}[\mathrm{Par}_{\mathcal{A}}]$ is more than a *set*, but is a set with *structure*. For deep feedforward neural networks, this structure manifests as the face poset of hyperplane arrangement, encapsulating both model expressivity and decision boundaries.

We studied the case when $\mathcal{A}$ is a Hopfield network, and the number of chambers, reduces to a problem of counting regions induced by hyperplane arrangements, and is explicitly given by Zaslavsky's theorem, [Sta07]. A simple corollary is:

Corollary 6.2 (Hillar, Lin)**.** *A Hopfield network of two nodes with asynchronous (or synchronous) updates cannot express XOR functions.*

We are currently focusing on parameter estimation: the process of using data to infer the values of unknown parameters within a model. Where are the critical points, and how is the dynamics of parameter estimation reflected in the parameter space? [Koh+22] has studied a similar story for linear convolutional networks. Our particular estimation method of interest in *minimum probability flow*, [SBD20], which has been used by joint author Chris Hillar in the case of Hopfield networks, [HMK14].

We summarize here where the future research would focus on:

(1) Understanding a suitable notion of "equivalence class of networks". What can we say about pairs $(\mathcal{A}, \mathcal{T}), (\mathcal{A}', \mathcal{T}')$ and the differences in their induced face posets, $\Sigma_{\mathcal{T}}[\mathrm{Par}_{\mathcal{A}}], \Sigma_{\mathcal{T}'}[\mathrm{Par}_{\mathcal{A}'}]$

    Similarly, can one describe distributions on the $\mathrm{Par}_{\mathcal{A}}$ that corresponds to networks with certain properties?[8]

(2) Extend our analysis of architectures *higher order networks*, which includes transformer networks, and simplicial hopfield network [BF23]. $\mathrm{Par}_{\mathcal{A}}$ decomposes into semi-algebraic sets, and one may approach with the theory of splines, [LLL24].

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)

We 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].

[8] A similar question was asked in [Mon+14].

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. 3).

[AL21] Anschütz, Johannes and Le Bras, Arthur-César. "A Fourier Transform for Banach-Colmez spaces". In: *arXiv e-prints*, arXiv:2111.11116 (Nov. 2021), arXiv:2111.11116. arXiv: 2111.11116 [math.AG] (cit. on p. 4).

[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).

[Bea83] Beauville, A. "Quelques remarques sur la transformation de Fourier dans l'anneau de Chow d'une variété abélienne". In: *Algebraic geometry (Tokyo/Kyoto, 1982)*. Vol. 1016. Lecture Notes in Math. Springer, Berlin, 1983, pp. 238–260. ISBN: 3-540-12685-6. URL: https://doi.org/10.1007/BFb0099965 (cit. on p. 5).

[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. 8, 9).

[Bha23] Bhatt, Bhargav. *Prismatic F-gauges*. 2023 (cit. on pp. 1, 6).

[Bro14] Brown, Francis. *Motivic periods and the projective line minus three points*. 2014. arXiv: 1407.5165 [math.NT]. URL: https://arxiv.org/abs/1407.5165 (cit. on pp. 1, 6).

[Bro17] Brown, Francis. "Notes on motivic periods". In: *Commun. Number Theory Phys.* 11.3 (2017), pp. 557–655. ISSN: 1931-4523,1931-4531. URL: https://doi.org/10.4310/CNTP.2017.v11.n3.a2 (cit. on pp. 2, 6).

[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 pp. 1, 4, 5).

[CD19] Cisinski, Denis-Charles and Déglise, Frédéric. *Triangulated categories of mixed motives*. Springer Monographs in Mathematics. Springer, Cham, [2019] ©2019, pp. xlii+406. ISBN: 978-3-030-33241-9; 978-3-030-33242-6. URL: https://doi.org/10.1007/978-3-030-33242-6 (cit. on p. 5).

[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. 9).

[Che73] Chen, Kuo-tsai. "Iterated integrals of differential forms and loop space homology". In: *Ann. of Math. (2)* 97 (1973), pp. 217–246. ISSN: 0003-486X. URL: https://doi.org/10.2307/1970846 (cit. on p. 6).

[Che76] Chen, Kuo Tsai. "Reduced bar constructions on de Rham complexes". In: *Algebra, topology, and category theory (a collection of papers in honor of Samuel Eilenberg)*. Academic Press, New York-London, 1976, pp. 19–32 (cit. on p. 6).

[Con15] Conrad, Brian. “Non-split reductive groups over **Z**”. In: *Autours des schémas en groupes. Vol. II.* Vol. 46. Panor. Synthèses. Soc. Math. France, Paris, 2015, pp. 193–253. ISBN: 978-2-85629-819-0. URL: https://doi.org/10.1017/CBO9781316092439 (cit. on p. 6).

[Cyb89] Cybenko, George V. “Approximation by superpositions of a sigmoidal function”. In: *Mathematics of Control, Signals and Systems* 2 (1989), pp. 303–314. URL: https://api.semanticscholar.org/CorpusID:3958369 (cit. on p. 8).

[Del89] Deligne, Par P. “Le groupe fondamental de la droite projective moins trois points”. In: *Galois Groups over* $\mathbb{Q}$ *Proceedings of a Workshop Held March 23–27, 1987.* Springer. 1989, pp. 79–297 (cit. on p. 6).

[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. 8).

[Dri22] Drinfeld, Vladimir. *A stacky approach to crystals.* 2022. arXiv: 1810.11853 [math.AG]. URL: https://arxiv.org/abs/1810.11853 (cit. on p. 6).

[Far20] Fargues, Laurent. “Simple connexité des fibres d’une application d’Abel-Jacobi et corps de classes local”. In: *Ann. Sci. Éc. Norm. Supér. (4)* 53.1 (2020), pp. 89–124. ISSN: 0012-9593,1873-2151. URL: https://doi.org/10.24033/asens.2418 (cit. on p. 4).

[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 pp. 1, 3, 5, 7).

[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, 3, 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, 3).

[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 pp. 1, 4).

[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. 3).

[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. 3).

[GGW18] Gan, Wee Teck, Gao, Fan, and Weissman, Martin H. “L-groups and the Langlands program for covering groups: a historical introduction”. In: 398. L-groups and the Langlands program for covering groups. 2018, pp. 1–31. ISBN: 978-2-85629-845-9 (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).

[GR14] Gaitsgory, Dennis and Rozenblyum, Nick. *Crystals and D-modules.* 2014. arXiv: 1111.2087 [math.AG]. URL: https://arxiv.org/abs/1111.2087 (cit. on pp. 1, 6).

[Hai01] Hain, Richard. *Iterated Integrals and Algebraic Cycles: Examples and Prospects.* 2001. arXiv: math/0109204 [math.AG]. URL: https://arxiv.org/abs/math/0109204 (cit. on p. 6).

[Han21] Hansen, David. "Moduli of local shtukas and Harris's conjecture". In: *Tunis. J. Math.* 3.4 (2021), pp. 749–799. ISSN: 2576-7658,2576-7666. URL: https://doi.org/10.2140/tunis.2021.3.749 (cit. on p. 4).

[Has+24] Hasan, Junaid, Hassan, Hazem, Lin, Milton, Manivel, Marcella, McBeath, Lily, and Moonen, Ben. *Integral aspects of Fourier duality for abelian varieties*. 2024. arXiv: 2407.06184 [math.AG]. URL: https://arxiv.org/abs/2407.06184 (cit. on p. 5).

[HMK14] Hillar, Christopher, Mehta, Ram, and Koepsell, Kilian. "A hopfield recurrent neural network trained on natural images performs state-of-the-art image compression". In: 2014, pp. 4092–4096 (cit. on p. 9).

[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 pp. 2, 8).

[HR20] Haines, Thomas J. and Richarz, Timo. "Smoothness of Schubert varieties in twisted affine Grassmannians". In: *Duke Mathematical Journal* 169.17 (Nov. 2020). ISSN: 0012-7094. URL: http://dx.doi.org/10.1215/00127094-2020-0025 (cit. on p. 4).

[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. 8).

[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, 3–5).

[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. 8).

[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. 9).

[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. 9).

[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. 2, 8, 9).

[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).

[Kim16] Kim, Minhyong. *Arithmetic Chern-Simons Theory I*. 2016. arXiv: 1510.05818 [math.NT]. URL: https://arxiv.org/abs/1510.05818 (cit. on p. 4).

[Koh+22] Kohn, Kathlén, Merkh, Thomas, Montúfar, Guido, and Trager, Matthew. *Geometry of Linear Convolutional Networks*. 2022. arXiv: 2108.01538 [cs.LG]. URL: https://arxiv.org/abs/2108.01538 (cit. on p. 9).

[KW07] Kapustin, Anton and Witten, Edward. *Electric-Magnetic Duality And The Geometric Langlands Program*. 2007. arXiv: hep-th/0604151 [hep-th]. URL: https://arxiv.org/abs/hep-th/0604151 (cit. on p. 4).

[Lev+19] Levy, Andrew, Konidaris, George, Platt, Robert, and Saenko, Kate. *Learning Multi-Level Hierarchies with Hindsight*. 2019. arXiv: 1712.00948 [cs.AI]. URL: https://arxiv.org/abs/1712.00948 (cit. on p. 2).

[LLL24] Lai, Zehua, Lim, Lek-Heng, and Liu, Yucong. *Attention is a smoothed cubic spline*. 2024. arXiv: 2408.09624 [cs.AI]. URL: https://arxiv.org/abs/2408.09624 (cit. on p. 9).

[Lur10] Lurie, Jacob. "Moduli problems for ring spectra". In: *Proceedings of the International Congress of Mathematicians. Volume II*. Hindustan Book Agency, New Delhi, 2010, pp. 1099–1125. ISBN: 978-81-85931-08-3; 978-981-4324-32-8; 981-4324-32-9 (cit. on pp. 2, 7).

[Lur18] Lurie, Jacob. "Spectral Algebraic Geometry". In: (2018) (cit. on p. 7).

[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 p. 2).

[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 pp. 2, 9).

[MR23] Mondal, Shubhodip and Reinecke, Emanuel. *Unipotent homotopy theory of schemes*. 2023. arXiv: 2302.10703 [math.AG] (cit. on pp. 1, 6).

[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. 9).

[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. 9).

[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 pp. 1, 3, 5).

[Pap07] Pappas, Georgios. "Integral Grothendieck–Riemann–Roch theorem". In: *Inventiones mathematicae* 170.3 (July 2007), pp. 455–481. ISSN: 1432-1297. URL: http://dx.doi.org/10.1007/s00222-007-0067-9 (cit. on p. 5).

[PZ24] Pappas, Georgios and Zhou, Rong. *On the smooth locus of affine Schubert varieties*. 2024. arXiv: 2312.14827 [math.AG]. URL: https://arxiv.org/abs/2312.14827 (cit. on p. 4).

[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. 8).

[RS20] Richarz, Timo and Scholbach, Jakob. "The intersection motive of the moduli stack of Shtuka". In: *Forum of Mathematics, Sigma* 8 (2020). ISSN: 2050-5094 (cit. on pp. 1, 5).

[SB18] Sutton, Richard S and Barto, Andrew G. *Reinforcement learning: An introduction*. MIT press, 2018 (cit. on p. 2).

[SBD20] Sohl-Dickstein, Jascha, Battaglino, Peter, and DeWeese, Michael R. *A new method for parameter estimation in probabilistic models: Minimum probability flow*. 2020. arXiv: 2007.09240 [cs.LG]. URL: https://arxiv.org/abs/2007.09240 (cit. on p. 9).

[She+24] Shehper, Ali, Medina-Mardones, Anibal M., Lewandowski, Bartłomiej, Gruen, Angus, Kucharski, Piotr, and Gukov, Sergei. *What makes math problems hard for reinforcement learning: a case study*. 2024. arXiv: 2408.15332 [cs.LG]. URL: https://arxiv.org/abs/2408.15332 (cit. on p. 2).

[Sta07] Stanley, Richard P. "An Introduction to Hyperplane Arrangements". In: 2007 (cit. on p. 8).

[Toë06] Toën, Bertrand. "Champs affines". In: *Selecta Math. (N.S.)* 12.1 (2006), pp. 39–135. ISSN: 1022-1824,1420-9020. URL: https://doi.org/10.1007/s00029-006-0019-z (cit. on pp. 1, 6).

[Vol01] Vologodsky, Vadim. *Hodge structure on the fundamental group and its application to p-adic integration*. 2001. arXiv: math/0108109 [math.AG] (cit. on pp. 2, 6).

[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 pp. 1, 4).

[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).