Novelty Assessment Report

Paper: Overparametrization bends the landscape: BBP transitions at initialization in simple Neural Networks

PDF URL: https://openreview.net/pdf?id=xDLE5n3x9Y

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2026-01-05

Abstract

High-dimensional non-convex loss landscapes play a central role in the theory of Machine Learning. Gaining insight into how these landscapes interact with gradient-based optimization methods, even in relatively simple models, can shed light on this enigmatic feature of neural networks. In this work, we will focus on a prototypical simple learning problem, which generalizes the Phase Retrieval inference problem by allowing the exploration of overparametrized settings. Using techniques from field theory, we analyze the spectrum of the Hessian at initialization and identify a Baik–Ben Arous–Péché (BBP) transition in the amount of data that separates regimes where the initialization is informative or uninformative about a planted signal of a teacher-student setup. Crucially, we demonstrate how overparameterization can "bend" the loss landscape, shifting the transition point, even reaching the information-theoretic weak-recovery threshold in the large overparameterization limit, while also altering its qualitative nature. We distinguish between continuous and discontinuous BBP transitions and support our analytical predictions with simulations, examining how they compare to the finite-N behavior. In the case of discontinuous BBP transitions strong finite-N corrections allow the retrieval of information at a signal-to-noise ratio (SNR) smaller than the predicted BBP transition. In these cases we provide estimates for a new lower SNR threshold that marks the point at which initialization becomes entirely uninformative.

Core Task Landscape

This paper addresses: **Hessian Spectrum Analysis at Initialization in Overparametrized Neural Networks**

A total of **16 papers** were analyzed and organized into a taxonomy with **9 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Theoretical Characterization of Hessian Spectral Properties**
- **Empirical Analysis of Hessian Structure and Eigenvalues**
- **Initialization Schemes and Their Impact**
- **Training Dynamics and Hessian Evolution**
- **Specialized Architectures and Reparameterizations**

Complete Taxonomy Tree

- Hessian Spectrum Analysis at Initialization in Overparametrized Neural Networks Survey Taxonomy
- Theoretical Characterization of Hessian Spectral Properties
  - Asymptotic Spectral Analysis and Random Matrix Theory (3 papers)
  - [1] Optimization and bayes: a trade-off for overparameterized neural networks (Z Hu, 2023) View paper
  - [3] Towards quantifying the hessian structure of neural networks (Zhang Yushun, 2025) View paper
  - [13] The asymptotic spectrum of the hessian of dnn throughout training (Jacot, 2019) View paper
  - Phase Transitions and Critical Phenomena at Initialization ★ (2 papers)
  - [0] Overparametrization bends the landscape: BBP transitions at initialization in simple Neural Networks (Anon et al., 2026) View paper
  - [11] Deconstructing the Goldilocks Zone of Neural Network Initialization (Vysogorets, 2024) View paper
- Empirical Analysis of Hessian Structure and Eigenvalues
  - Bulk and Edge Eigenvalue Distribution (2 papers)
  - [4] Eigenvalues of the Hessian in Deep Learning: Singularity and Beyond (Sagun, 2022) View paper
  - [8] Neglected hessian component explains mysteries in sharpness regularization (Atish Agarwala, 2024) View paper
  - Structural Properties and Block-Diagonal Patterns (1 papers)
  - [7] Statistical mechanics of deep learning (Sompolinsky, 2020) View paper
- Initialization Schemes and Their Impact
  - Variance-Based and Spectral Norm Initialization (2 papers)
  - [2] Revisiting weight initialization of deep neural networks (Skorski, 2021) View paper
  - [5] Revisiting Initialization of Neural Networks (Skorski, 2022) View paper
  - Initialization Effects on Generalization and Implicit Bias (2 papers)
  - [6] How much does initialization affect generalization? (S Ramasinghe, 2023) View paper
  - [16] How Width Scaling Affects Neural Networks: Generalization, Optimal Hyperparameters, Feature Learning and Beyond (Haas, n.d.) View paper
- Training Dynamics and Hessian Evolution
  - Saddle Point Escape and Early Training Dynamics (1 papers)
  - [10] Saddle-To-Saddle Dynamics in Deep ReLU Networks: Low-Rank Bias in the First Saddle Escape (Simon, 2025) View paper
  - Time-Dependent Hessian and Spectral Transitions During Training (1 papers)

- ◦ [14] The role of the time-dependent Hessian in high-dimensional optimization (Tony Bonnaire, 2025) View paper
- Specialized Architectures and Reparameterizations (3 papers)
  - ◦ [9] Shallow univariate ReLu networks as splines: initialization, loss surface, hessian, and gradient flow dynamics (Justin Sahs, 2022) View paper
  - ◦ [12] Shallow Univariate ReLu Networks as Splines: Initialization, Loss Surface, Hessian, & Gradient Flow Dynamics (Sahs, 2022) View paper

Narrative

Core task: Hessian spectrum analysis at initialization in overparametrized neural networks. The field examines how the eigenvalue distribution of the loss Hessian at initialization shapes subsequent training and generalization. The taxonomy organizes this landscape into several main branches. Theoretical Characterization of Hessian Spectral Properties investigates phase transitions, critical phenomena, and asymptotic spectral laws that emerge as network width or depth grows, often drawing on random matrix theory and statistical mechanics perspectives. Empirical Analysis of Hessian Structure and Eigenvalues focuses on measuring and quantifying eigenvalue distributions, bulk versus outlier structure, and the role of neglected components in real networks. Initialization Schemes and Their Impact studies how different weight-scaling strategies (e.g., Xavier, He, or novel parameterizations) alter the Hessian spectrum and downstream optimization. Training Dynamics and Hessian Evolution tracks how eigenvalues shift during gradient descent, revealing saddle-point structure and time-dependent spectral changes. Finally, Specialized Architectures and Reparameterizations explores how architectural choices (residual connections, normalization layers, or custom parameterizations) modify the Hessian at initialization.

A particularly active line of work examines phase transitions and critical regimes where small changes in initialization scale trigger qualitative shifts in the Hessian spectrum, as seen in BBP Transitions Initialization[0] and Goldilocks Zone Initialization[11], which identify narrow windows of initialization variance that balance trainability and feature learning. These studies contrast with broader empirical investigations like Quantifying Hessian Structure[3] and Neglected Hessian Component[8], which document how bulk eigenvalue distributions and often-ignored spectral components influence optimization trajectories. BBP Transitions Initialization[0] sits squarely within the theoretical characterization branch, emphasizing critical phenomena at initialization and connecting spectral properties to subsequent training phases. Its focus on phase boundaries complements works like Asymptotic Hessian Spectrum[13], which derives limiting spectral densities, and Goldilocks Zone Initialization[11], which empirically validates the existence of optimal initialization regimes. Together, these efforts reveal that initialization is not merely a practical detail but a window into the geometry and trainability of overparametrized models.

Related Works in Same Category

The following **1 sibling papers** share the same taxonomy leaf node with the original paper:

1. Deconstructing the Goldilocks Zone of Neural Network Initialization

Authors: Vysogorets, Artem, Dawid, Anna, Artem Vysogorets, et al. (9 authors total) | **Year/Venue**: 2024 • International Conference on Machine Learning | **URL**: View paper

Abstract

The second-order properties of the training loss have a massive impact on the optimization dynamics of deep learning models. Fort&Scherlis (2019) discovered that a large excess of positive curvature and local convexity of the loss Hessian is associated with highly trainable initial points located in a region coined the"Goldilocks zone". Only a handful of subsequent studies touched upon this relationship, so it remains largely unexplained. In this paper, we present a rigorous and comprehensive an...

Relationship Analysis

Both papers belong to the category of phase transitions and critical phenomena at initialization, examining how Hessian spectral properties change as parameters vary. The original paper focuses on BBP transitions in overparametrized teacher-student networks, analyzing how the amount of data (SNR) affects when the Hessian spectrum becomes informative about planted signals, while the candidate paper investigates the "Goldilocks zone" phenomenon, examining how initialization norm affects the excess of positive curvature in the Hessian and its relationship to trainability. The key difference is that the original paper studies phase transitions in signal recovery as a function of dataset size in a specific inference problem, whereas the candidate paper analyzes the relationship between initialization scale and local convexity properties that enable effective training.

Contributions Analysis

Overall novelty summary. The paper contributes a field-theoretic analysis of Hessian spectra at initialization in overparametrized neural networks, identifying Baik–Ben Arous–Péché (BBP) transitions that separate informative from uninformative initialization regimes. It resides in the 'Phase Transitions and Critical Phenomena at Initialization' leaf, which contains only two papers total. This sparse population suggests the paper addresses a relatively specialized research direction within the broader Hessian analysis landscape, focusing on critical phenomena rather than general spectral characterization or empirical measurement.

The taxonomy tree reveals that the paper's immediate parent branch, 'Theoretical Characterization of Hessian Spectral Properties', contains a sibling leaf on 'Asymptotic Spectral Analysis and Random Matrix Theory' with three papers. Neighboring branches include 'Empirical Analysis of Hessian Structure' (three papers across two leaves) and 'Initialization Schemes and Their Impact' (four papers). The paper's use of field theory and random matrix techniques connects it to asymptotic spectral work, while its focus on overparameterization and information-theoretic thresholds distinguishes it from purely empirical eigenvalue distribution studies or initialization scheme proposals.

Among thirty candidates examined, none were found to clearly refute any of the three main contributions. Contribution A (BBP transitions in overparametrized networks) examined ten candidates with zero refutable matches; Contribution B (continuous versus discontinuous transitions) and Contribution C (weak-recovery threshold via infinite overparametrization) each examined ten candidates with identical outcomes. This suggests that within the limited search scope, the specific combination of BBP transition analysis, overparameterization effects, and information-theoretic threshold characterization appears relatively unexplored, though the absence of refutations does not guarantee exhaustive novelty.

Given the sparse taxonomy leaf (two papers) and zero refutations across thirty candidates, the work appears to occupy a distinct niche within Hessian initialization theory. However, the limited search scope means potentially relevant work in statistical physics, phase retrieval, or teacher-student frameworks outside the top-thirty semantic matches may not have been captured. The analysis reflects novelty within the examined literature but cannot rule out overlooked connections in adjacent fields.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Analysis of BBP transitions in overparametrized neural networks at initialization

Description: The authors apply field-theoretic techniques to study the Hessian spectrum at initialization in a teacher-student setup with two-layer networks. They characterize the BBP transition that determines when random initialization contains information about the teacher signal, extending this analysis to overparametrized settings beyond standard phase retrieval.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Revisiting Initialization of Neural Networks

URL: View paper

Brief Assessment

Revisiting Network Initialization[5] focuses on Hessian-based weight initialization schemes to control global curvature for training stability, not on analyzing BBP transitions in teacher-student setups or characterizing when random initialization contains information about planted signals.

2. The Challenges of the Nonlinear Regime for Physics-Informed Neural Networks

URL: View paper

Brief Assessment

Nonlinear Regime PINN[17] focuses on physics-informed neural networks for solving PDEs and analyzes the Hessian spectrum in the context of NTK theory for linear vs. nonlinear PDEs. It does not address BBP transitions in teacher-student setups or phase retrieval problems, which are the core focus of the original contribution.

3. Eigenvalues of the Hessian in Deep Learning: Singularity and Beyond

URL: View paper

Brief Assessment

Hessian Eigenvalues Singularity[4] focuses on the eigenvalue distribution of the Hessian after training in fully-connected networks, examining singularity and bulk concentration around zero. It does not analyze BBP transitions, teacher-student setups, or information-theoretic recovery thresholds at initialization.

4. Towards quantifying the hessian structure of neural networks

URL: View paper

Brief Assessment

Quantifying Hessian Structure[3] focuses on the block-diagonal structure of Hessian matrices at initialization for classification tasks, not on BBP transitions in teacher-student setups or phase retrieval problems. The candidate analyzes how the number of classes affects Hessian structure, while the original paper studies BBP transitions that determine when random initialization contains information about the teacher signal in overparametrized settings.

5. Shallow Univariate ReLu Networks as Splines: Initialization, Loss Surface, Hessian, & Gradient Flow Dynamics

URL: View paper

Brief Assessment

Shallow ReLU Splines[12] focuses on loss surface structure and implicit regularization in shallow univariate ReLU networks using spline reparametrization, not on BBP transitions or Hessian spectrum analysis in teacher-student setups.

6. Shallow univariate ReLu networks as splines: initialization, loss surface, hessian, and gradient flow dynamics

URL: View paper

Brief Assessment

Shallow ReLU Splines[9] focuses on univariate shallow ReLU networks analyzed through a spline lens, examining initialization distributions, loss surface geometry, and implicit regularization. It does not address Hessian spectrum analysis at initialization in teacher-student setups or BBP transitions in overparametrized networks.

7. The asymptotic spectrum of the hessian of dnn throughout training

URL: View paper

Brief Assessment

Asymptotic Hessian Spectrum[13] focuses on the Hessian spectrum throughout training in the infinite-width limit using the Neural Tangent Kernel framework, not specifically on BBP transitions at initialization in teacher-student setups with field-theoretic techniques.

8. Deconstructing the Goldilocks Zone of Neural Network Initialization

URL: View paper

Brief Assessment

Goldilocks Zone Initialization[11] focuses on second-order properties and positive curvature at initialization for trainability, not on BBP transitions or Hessian spectrum analysis in teacher-student setups.

9. Fishing For Cheap And Efficient Pruners At Initialization

URL: View paper

Brief Assessment

Pruners At Initialization[19] focuses on neural network pruning using Fisher Information Matrix diagonal approximations at initialization, not on analyzing Hessian spectrum BBP transitions in teacher-student setups or phase retrieval problems.

10. Vanishing Curvature and the Power of Adaptive Methods in Randomly Initialized Deep Networks

URL: View paper

Brief Assessment

Vanishing Curvature Adaptive[18] focuses on vanishing gradients and Hessian eigenspectra in deep randomly initialized networks, analyzing how depth affects curvature. The original paper studies BBP transitions in a teacher-student setup with field-theoretic techniques to characterize when initialization contains information about the teacher signal—a fundamentally different problem.

Contribution 2: Characterization of continuous versus discontinuous BBP transitions under overparametrization

Description: The authors identify and distinguish two qualitatively different types of BBP transitions (continuous and discontinuous) that arise depending on overparametrization level and loss normalization. They show that higher overparametrization systematically leads to discontinuous transitions with strong finite-size effects.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Learning through atypical phase transitions in overparameterized neural networks

URL: View paper

Brief Assessment

[Final Audit Failure] The model insisted on a refutation claim but failed to provide verifiable evidence after multiple retries. Marked as cannot_refute for safety. Please manually verify the candidate text.

---

2. Theory of overparametrization in quantum neural networks

URL: View paper

Brief Assessment

Quantum Neural Networks[32] studies overparametrization in quantum neural networks and phase transitions in quantum Fisher information matrix ranks, not classical neural networks with continuous/discontinuous BBP transitions in loss landscapes.

---

3. Understanding pathologies of deep heteroskedastic regression

URL: View paper

Brief Assessment

Heteroskedastic Regression Pathologies[37] studies phase transitions in heteroskedastic regression models with regularization parameters, not BBP transitions in overparametrized neural networks for signal recovery tasks as in the original paper.

---

4. Memorizing without overfitting: Bias, variance, and interpolation in overparameterized models

URL: View paper

Brief Assessment

Memorizing Without Overfitting[30] focuses on bias-variance decomposition in overparameterized models and phase transitions in training/test error, not on characterizing continuous versus discontinuous BBP spectral transitions in loss landscape Hessians.

---

5. Optimal generalisation and learning transition in extensive-width shallow neural networks near interpolation

URL: View paper

Brief Assessment

Optimal Generalisation Interpolation[34] studies learning transitions in neural networks near interpolation but focuses on universal versus specialisation phases rather than continuous/discontinuous BBP transitions in overparametrized settings.

---

6. Bias-variance decomposition of overparameterized regression with random linear features

URL: View paper

Brief Assessment

Bias Variance Random Features[38] focuses on bias-variance decomposition in overparameterized regression with random linear features, analyzing three distinct regimes separated by phase transitions. While it discusses phase transitions in the context of interpolation thresholds, it does not specifically characterize continuous versus discontinuous BBP transitions as qualitatively different phenomena arising from overparametrization levels and loss normalization as described in the original paper.

---

7. Hidden progress in deep learning: Sgd learns parities near the computational limit

URL: View paper

Brief Assessment

SGD Learns Parities[35] focuses on learning sparse parities with neural networks and analyzes phase transitions in training curves, not on characterizing continuous versus discontinuous BBP transitions in overparametrized settings as studied in the original paper.

---

8. Neural models for prediction of spatially patterned phase transitions: methods and challenges

URL: View paper

Brief Assessment

Spatially Patterned Transitions[33] focuses on neural network models for predicting spatially patterned phase transitions in dryland vegetation ecosystems, not on BBP transitions in overparametrized neural network learning landscapes. The paper addresses ecological regime shifts and early warning signals, which is a fundamentally different domain from the original paper's analysis of loss landscape geometry in machine learning.

---

9. A jamming transition from under-to over-parametrization affects generalization in deep learning

URL: View paper

Brief Assessment

Jamming Transition Parametrization[36] focuses on a jamming transition between under- and over-parametrized regimes affecting generalization, not on characterizing continuous versus discontinuous BBP transitions in spectral properties of the Hessian at initialization.

---

10. Overparameterized relu neural networks learn the simplest model: Neural isometry and phase transitions

URL: View paper

Brief Assessment

Simplest Model Neural Isometry[25] focuses on phase transitions in recovering planted models via ReLU networks using convex optimization and sparse recovery perspectives. The candidate does not address BBP transitions in loss landscape Hessian spectra at initialization or their qualitative nature (continuous vs. discontinuous) under overparametrization.

Contribution 3: Demonstration that infinite overparametrization achieves information-theoretic weak-recovery threshold

Description: The authors prove that in the limit of infinite overparametrization, the BBP transition threshold converges to the information-theoretic weak-recovery threshold. This shows that spectral analysis of the Hessian at initialization can match optimal recovery performance through overparametrization alone.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. A validation approach to over-parameterized matrix and image recovery

URL: View paper

Brief Assessment

Overparameterized Matrix Recovery[20] focuses on matrix recovery from noisy linear measurements using gradient descent with early stopping, not on spectral analysis of Hessians at initialization or BBP transitions in neural network loss landscapes.

2. Overparameterization from computational constraints

URL: View paper

Brief Assessment

[Final Audit Failure] The model insisted on a refutation claim but failed to provide verifiable evidence after multiple retries. Marked as cannot_refute for safety. Please manually verify the candidate text.

3. Flat minima generalize for low-rank matrix recovery

URL: View paper

Brief Assessment

Flat Minima Matrix Recovery[21] focuses on low-rank matrix factorization problems (matrix sensing, completion, robust PCA) and analyzes flat minima via Hessian trace. The original paper studies BBP transitions in neural networks with quadratic activations at initialization, examining how overparametrization affects spectral properties of the Hessian. These are fundamentally different problem settings and analytical approaches.

4. Can neural networks achieve optimal computational-statistical tradeoff? an analysis on single-index model

URL: View paper

Brief Assessment

Single Index Model[23] focuses on gradient-based training of neural networks for Gaussian single-index models with sample complexity O(d^(s*/2)), not on spectral analysis of Hessians at initialization for phase retrieval problems.

5. Information-theoretic reduction of deep neural networks to linear models in the overparametrized proportional regime

URL: View paper

Brief Assessment

[Final Audit Failure] The model insisted on a refutation claim but failed to provide verifiable evidence after multiple retries. Marked as cannot_refute for safety. Please manually verify the candidate text.

6. Local Linear Recovery Guarantee of Deep Neural Networks at Overparameterization

URL: View paper

Brief Assessment

Local Linear Recovery[27] focuses on local linear recovery guarantees for DNNs through model rank analysis, not on spectral analysis of Hessians at initialization or BBP transitions in overparameterized neural networks.

7. Overparameterized relu neural networks learn the simplest models: Neural isometry and exact recovery

URL: View paper

Brief Assessment

The candidate paper (Neural Isometry Exact Recovery[29]) focuses on exact recovery of planted neurons in two-layer ReLU networks through convex optimization and sparse recovery perspectives, not on BBP transitions or Hessian spectral analysis at initialization. The original paper studies how overparameterization affects the BBP transition in the Hessian spectrum at initialization for teacher-student setups, while the candidate examines recovery conditions for neural networks with arbitrary parameters through neural isometry conditions.

8. Bayes-optimal learning of an extensive-width neural network from quadratically many samples

URL: View paper

Brief Assessment

[Final Audit Failure] The model insisted on a refutation claim but failed to provide verifiable evidence after multiple retries. Marked as cannot_refute for safety. Please manually verify the candidate text.

9. Investigating over-parameterized randomized graph networks

URL: View paper

Brief Assessment

Overparameterized Graph Networks[22] investigates graph neural networks using algorithmic stability and information theory measures, not overparametrization effects on weak-recovery thresholds in teacher-student setups with spectral methods.

10. Overparameterized relu neural networks learn the simplest model: Neural isometry and phase transitions

URL: View paper

Brief Assessment

While Simplest Model Neural Isometry[25] discusses recovery thresholds and phase transitions in overparameterized networks, it analyzes recovery of planted models through convex reformulations rather than BBP transitions in Hessian spectra. The candidate's focus on spectral methods and recovery conditions differs from analyzing Hessian eigenvalue transitions at initialization.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Overparametrization bends the landscape: BBP transitions at initialization in simple Neural Networks View paper
- [1] Optimization and bayes: a trade-off for overparameterized neural networks View paper
- [2] Revisiting weight initialization of deep neural networks View paper
- [3] Towards quantifying the hessian structure of neural networks View paper
- [4] Eigenvalues of the Hessian in Deep Learning: Singularity and Beyond View paper
- [5] Revisiting Initialization of Neural Networks View paper
- [6] How much does initialization affect generalization? View paper
- [7] Statistical mechanics of deep learning View paper

- [8] Neglected hessian component explains mysteries in sharpness regularization View paper
- [9] Shallow univariate ReLu networks as splines: initialization, loss surface, hessian, and gradient flow dynamics View paper
- [10] Saddle-To-Saddle Dynamics in Deep ReLU Networks: Low-Rank Bias in the First Saddle Escape View paper
- [11] Deconstructing the Goldilocks Zone of Neural Network Initialization View paper
- [12] Shallow Univariate ReLu Networks as Splines: Initialization, Loss Surface, Hessian, & Gradient Flow Dynamics View paper
- [13] The asymptotic spectrum of the hessian of dnn throughout training View paper
- [14] The role of the time-dependent Hessian in high-dimensional optimization View paper
- [15] Shallow Univariate ReLu Networks as Splines: Initialization, Loss Surface, Hessian, & Gradient Flow Dynamics View paper
- [16] How Width Scaling Affects Neural Networks: Generalization, Optimal Hyperparameters, Feature Learning and Beyond View paper
- [17] The Challenges of the Nonlinear Regime for Physics-Informed Neural Networks View paper
- [18] Vanishing Curvature and the Power of Adaptive Methods in Randomly Initialized Deep Networks View paper
- [19] Fishing For Cheap And Efficient Pruners At Initialization View paper
- [20] A validation approach to over-parameterized matrix and image recovery View paper
- [21] Flat minima generalize for low-rank matrix recovery View paper
- [22] Investigating over-parameterized randomized graph networks View paper
- [23] Can neural networks achieve optimal computational-statistical tradeoff? an analysis on single-index model View paper
- [24] Information-theoretic reduction of deep neural networks to linear models in the overparametrized proportional regime View paper
- [25] Overparameterized relu neural networks learn the simplest model: Neural isometry and phase transitions View paper
- [26] Overparameterization from computational constraints View paper
- [27] Local Linear Recovery Guarantee of Deep Neural Networks at Overparameterization View paper
- [28] Bayes-optimal learning of an extensive-width neural network from quadratically many samples View paper
- [29] Overparameterized relu neural networks learn the simplest models: Neural isometry and exact recovery View paper
- [30] Memorizing without overfitting: Bias, variance, and interpolation in overparameterized models View paper
- [31] Learning through atypical phase transitions in overparameterized neural networks View paper
- [32] Theory of overparametrization in quantum neural networks View paper
- [33] Neural models for prediction of spatially patterned phase transitions: methods and challenges View paper
- [34] Optimal generalisation and learning transition in extensive-width shallow neural networks near interpolation View paper
- [35] Hidden progress in deep learning: Sgd learns parities near the computational limit View paper
- [36] A jamming transition from under-to over-parametrization affects generalization in deep learning View paper
- [37] Understanding pathologies of deep heteroskedastic regression View paper
- [38] Bias-variance decomposition of overparameterized regression with random linear features View paper