# Intruder Dimensions: A Random Matrix Theory Perspective on LoRA versus Full Fine-Tuning

Borna Khodabandeh[*] Sepehr Heidari Advary[*]
Department of Electrical Engineering
Sharif University of Technology
Emails: {borna710kh, sepehr.heid81}@gmail.com

*Abstract*—**Fine-tuning large pre-trained models is a critical step in adapting them to downstream tasks. LoRA (Low-Rank Adaptation) has emerged as a parameter-efficient alternative to full fine-tuning by introducing low-rank updates to weight matrices. However, despite achieving comparable task performance, LoRA and full fine-tuning yield fundamentally different spectral properties. In this work, we leverage Random Matrix Theory (RMT) to analyze how LoRA affects the singular value and singular vector structure of pre-trained weight matrices. We demonstrate that LoRA introduces intruder dimensions—new singular vectors that misalign with the pre-trained spectral basis—via the Baik-Ben Arous-Péché (BBP) phase transition. These spectral outliers degrade generalization and stability, especially in sequential learning settings. We further analyze mitigation strategies, including rank-stabilized LoRA and spectral fine-tuning, which preserve the pre-trained spectral structure while maintaining parameter efficiency. Our findings provide a rigorous theoretical foundation for understanding parameter-efficient fine-tuning methods and inform strategies for designing more robust adaptation techniques.** ***High-Dimentional-Probability-Analysis Course Project Winter 2025.***

## I. Introduction

In recent years, deep neural networks have achieved state-of-the-art performance on a wide variety of tasks, ranging from natural language processing to computer vision. A significant breakthrough in this domain has been the development of large pre-trained models that can be fine-tuned for specific downstream applications. Fine-tuning these models is a critical step in adapting them to new tasks. Traditional full fine-tuning updates every parameter of the pre-trained model, which is computationally expensive and can lead to overfitting.

An alternative approach, known as Low-Rank Adaptation (LoRA), has emerged as a parameter-efficient method. LoRA operates by injecting low-rank updates into the weight matrices of the pre-trained network. Despite empirical evidence that LoRA achieves performance comparable to full fine-tuning on target tasks, recent studies indicate that the two approaches yield substantially different spectral properties of the model's weight matrices.

In this project, we employ Random Matrix Theory (RMT) to rigorously analyze these spectral differences. We focus particularly on the phenomenon of *intruder dimensions*—new singular vectors that are introduced by the low-rank update and are misaligned with the original pre-trained spectral basis. These intruder dimensions are theorized to emerge via the Baik-Ben Arous-Péché (BBP) phase transition, a well-known effect in RMT when a low-rank perturbation is added to a random matrix. The presence of these outlier singular values can degrade the generalization and stability of the adapted model, especially when the model is subject to sequential learning tasks.

The remainder of this report is organized as follows. In Section II, we provide background on Random Matrix Theory and related work on spectral analysis in machine learning. Section III details our modeling of the pre-trained weight matrix and the fine-tuning processes (both full fine-tuning and LoRA). Section IV presents a rigorous RMT analysis, including derivations based on the resolvent method and the BBP transition. In Section V, we discuss strategies to mitigate the adverse effects of intruder dimensions, including increasing the rank and rank stabilization via spectral fine-tuning mechanisms. Section **??** describes simulation experiments that support our theoretical analysis. We conclude in Section VI with a discussion of the implications of our findings and directions for future work.

## II. Background and Related Work

Random Matrix Theory (RMT) studies the statistical properties of matrices whose entries are random variables. It has found applications in diverse fields such as quantum information theory, statistical physics, finance, and machine learning. Key areas of focus in RMT include:

- The joint distribution of eigenvalues.
- The behavior and distribution of eigenvectors.
- The asymptotic empirical spectral distribution (ESD) of large random matrices.
- The distribution of spacings between eigenvalues.
- Analysis of spiked matrix models, where a low-rank perturbation is added to a random matrix.
- Perturbation analysis and phase transitions in spectral properties.

### *A. Classical Ensembles and Universality*

Two of the most fundamental ensembles studied in RMT are the **Wigner ensemble** and the **Wishart ensemble**.

*1) Wigner Matrices and the Semicircle Law:* A Wigner matrix is a symmetric (or Hermitian) matrix $\mathbf{W} \in R^{n\times n}$ (or $C^{n\times n}$) whose entries $W_{ij}$ are independent (up to the symmetry constraint) random variables with mean zero and variance $\sigma^2/n$ for $i \neq j$ (and possibly a different variance on the diagonal). In the large $n$ limit, the empirical spectral distribution of a Wigner matrix converges to the **semicircle law**:

$$\rho_{\text{sc}}(x) = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2}\,\mathbf{1}_{\{|x|\leq 2\sigma\}},$$

where $\mathbf{1}_{\{|x|\leq 2\sigma\}}$ denotes the indicator function on the interval $[-2\sigma, 2\sigma]$.

*2) Wishart Matrices and the Marchenko–Pastur Law:* The Wishart ensemble arises in statistics and represents sample covariance matrices. Let $\mathbf{X} \in R^{p\times n}$ be a matrix with independent entries satisfying $E[X_{ij}] = 0$ and $\text{Var}(X_{ij}) = \sigma^2$. The sample covariance matrix is given by:

$$\mathbf{W} = \frac{1}{n}\mathbf{X}\mathbf{X}^\top.$$

In the high-dimensional limit, where $p, n \to \infty$ with $c = p/n$ fixed, the empirical spectral distribution of $\mathbf{W}$ converges to the **Marchenko–Pastur (MP) law**:

$$\rho_{\text{MP}}(x) = \frac{1}{2\pi\sigma^2 c\, x}\sqrt{(x-\lambda_-)(\lambda_+ - x)}\,\mathbf{1}_{\{x\in[\lambda_-,\lambda_+]\}},$$

where

$$\lambda_\pm = \sigma^2\left(1 \pm \sqrt{c}\right)^2.$$

*3) Universality:* One of the most remarkable aspects of RMT is the concept of **universality**. This principle asserts that many spectral properties, such as the semicircle law and the MP law, do not depend on the precise distribution of the matrix entries but only on certain moment conditions (e.g., finite variance). Hence, even if the entries of the matrix are non-Gaussian, the same limiting spectral distributions are observed under broad conditions.

In the context of deep learning, recent studies have begun to explore the spectral properties of weight matrices in neural networks. It has been observed that full fine-tuning tends to preserve the spectral structure inherited from pre-training, while LoRA introduces distinct outlier singular values. Such spectral differences have implications for model generalization and stability.

### B. The Baik–Ben Arous–Péché (BBP) Phase Transition

A key phenomenon in the study of spiked random matrix models is the *BBP phase transition* [1]–[3], [7]. Consider a random matrix whose spectral measure follows, for example, the Marchenko–Pastur law. When a low-rank (typically rank-one) perturbation is added to such a matrix, an additional eigenvalue may separate from the bulk spectrum if the perturbation (or "spike") exceeds a critical threshold.

More formally, let the perturbed spectral measure be denoted by $\mu_\theta(dx)$ when a spike of strength $\theta$ is introduced. Under appropriate normalization, one can express this measure as (assuming $\sigma^2 = 1$)

$$\mu_\theta(dx) = \frac{\sqrt{4-x^2}}{2\pi\big(\theta^2 + 1 - \theta x\big)}\mathbf{1}_{\{|x|<2\}}\,dx + \mathbf{1}_{\{|\theta|\geq\sqrt{\lambda_+}\}}\Big(1 - \frac{1}{\theta^2}\Big)\delta_{\theta+\frac{1}{\theta}}(dx). \quad (1)$$

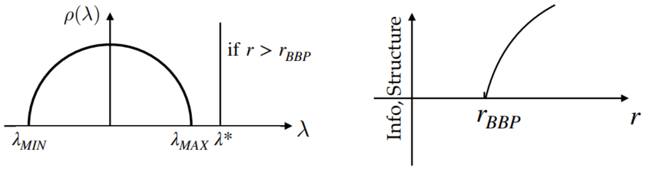

Fig. 1. General mechanism for emergence of information (or specific underlying structure)

In this expression:

- The first term represents the deformed bulk density, where the indicator $\mathbf{1}_{\{|x|<2\}}$ restricts the support to the interval $[-2, 2]$.
- The second term is a Dirac mass at the outlier eigenvalue $\theta + \frac{1}{\theta}$, which appears only when $|\theta|$ exceeds the critical threshold $\sqrt{\lambda_+}$ (with $\lambda_+$ being the upper edge of the bulk spectrum).

This formulation captures the essence of the BBP transition: when the spike $\theta$ is below the threshold, the perturbation is absorbed into the bulk spectrum; once $\theta$ exceeds $\sqrt{\lambda_+}$, an outlier (or *intruder dimension*) emerges. Such behavior is crucial for understanding how low-rank updates—such as those employed in LoRA—alter the spectral properties of pre-trained weight matrices.

The appearance of these outliers not only alters the spectral measure but also has significant implications for model generalization and stability, as the new directions introduced by the spike may not align with the original, pre-trained structure.

## III. Methodology

In this section, we describe our approach to modeling the weight matrices and fine-tuning procedures, and we outline the theoretical tools from RMT that are used in our analysis.

### A. Modeling the Pre-Trained Weight Matrix

We model the pre-trained weight matrix, $\mathbf{W}_{\text{pre}} \in R^{N\times N}$, as a random matrix with independent, identically distributed (i.i.d.) Gaussian entries (actually not needed because of Universality):

$$W_{ij} \sim \mathcal{N}\left(0, \frac{\sigma^2}{N}\right).$$

For large $N$, the singular values (or eigenvalues, if the matrix is symmetric) of $\mathbf{W}_{\text{pre}}$ are known to follow the Marchenko-Pastur distribution $\rho_{\text{MP}}(\lambda)$. This distribution is given by:

$$\rho_{\text{MP}}(\lambda) = \frac{1}{2\pi\sigma^2}\frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda} + \max\Big(1 - \tfrac{1}{c}, 0\Big)\delta_0 \quad \lambda \in [\lambda_-, \lambda_+], \quad (2)$$

where

$$\lambda_\pm = \sigma^2(1 \pm \sqrt{c})^2,$$

and $c$ is the limiting aspect ratio of the matrix dimensions.

### *B. Fine-Tuning Procedures*

We consider two fine-tuning methods:

1) **Full Fine-Tuning (FT):** A dense perturbation is applied to every element of $\mathbf{W}_{\text{pre}}$, resulting in the fine-tuned weight matrix:

$$\mathbf{W}_{\text{FT}} = \mathbf{W}_{\text{pre}} + \Delta\mathbf{W}_{\text{FT}},$$

where $\Delta\mathbf{W}_{\text{FT}}$ is a matrix with small perturbations.

2) **LoRA (Low-Rank Adaptation):** A low-rank update is introduced in the form:

$$\mathbf{W}_{\text{LoRA}} = \mathbf{W}_{\text{pre}} + \Delta\mathbf{W}_{\text{LoRA}}, \quad \Delta\mathbf{W}_{\text{LoRA}} = \mathbf{B}\mathbf{A}^\top,$$

where $\mathbf{B}, \mathbf{A} \in R^{N\times r}$ with $r \ll N$.

The critical difference between the two methods lies in the structure of the perturbation: full fine-tuning applies a dense, nearly isotropic update, whereas LoRA's update is confined to a low-dimensional subspace.

### *C. Full Fine-Tuning and Perturbative Corrections*

In full fine-tuning, the pre-trained weight matrix $\mathbf{W}_{\text{pre}}$ is updated via a dense perturbation:

$$\mathbf{W}_{\text{FT}} = \mathbf{W}_{\text{pre}} + \Delta\mathbf{W}_{\text{FT}},$$

where $\Delta\mathbf{W}_{\text{FT}}$ is a matrix with small entries applied uniformly across all parameters.

Using classical perturbation theory, the first-order correction to an eigenvalue $\lambda_i$ of $\mathbf{W}_{\text{pre}}$ is given by:

$$\lambda_i^{(1)} = v_i^\top \Delta\mathbf{W}_{\text{FT}}\, v_i,$$

where $v_i$ is the eigenvector corresponding to $\lambda_i$.

Similarly, the first-order correction to the eigenvector $v_i$ is expressed as:

$$v_i^{(1)} = \sum_{j\neq i} \frac{v_j^\top \Delta\mathbf{W}_{\text{FT}}\, v_i}{\lambda_i - \lambda_j}\, v_j.$$

Because $\Delta\mathbf{W}_{\text{FT}}$ is small and dense, these corrections are minor. In particular, the eigenvector correction $v_i^{(1)}$ induces only a small rotation of the original eigenbasis $\{v_i\}$, thereby preserving the overall spectral structure of the pre-trained model.

This small rotation ensures that the intrinsic features learned during pre-training remain largely intact, contributing to the robust generalization observed in full fine-tuning.

## IV. Spectral Analysis via Random Matrix Theory

This section details the RMT-based analysis of the spectral properties of $\mathbf{W}_{\text{pre}} + \Delta\mathbf{W}$, with emphasis on the effects of low-rank perturbations.

### *A. The Resolvent Method and Perturbation Analysis*

A central tool in RMT is the resolvent (or Green's function) of a matrix, defined as:

$$G(z) = (zI - \mathbf{W}_{\text{pre}})^{-1}, \quad z \in C \setminus \text{spec}(\mathbf{W}_{\text{pre}}).$$

The resolvent encapsulates information about the spectrum of $\mathbf{W}_{\text{pre}}$. In particular, the Stieltjes transform of the empirical spectral measure $\mu_{\mathbf{W}_{\text{pre}}}$ is given by:

$$m(z) = \frac{1}{N}\operatorname{tr} G(z) = \int \frac{d\mu_{\mathbf{W}_{\text{pre}}}(\lambda)}{z - \lambda}.$$

*1) Sherman–Morrison Formula:* When a rank-1 update is applied, we consider:

$$\Delta\mathbf{W} = \theta\, uv^\top, \quad \|u\| = \|v\| = 1, \quad \theta > 0.$$

Then, the perturbed matrix is:

$$\mathbf{M} = \mathbf{W}_{\text{pre}} + \theta\, uv^\top.$$

The resolvent of $\mathbf{M}$ can be expressed using the Sherman–Morrison formula:

$$\begin{aligned} G_{\mathbf{M}}(z) = (zI - \mathbf{M})^{-1} &= \left(zI - \mathbf{W}_{\text{pre}} - \theta\, uv^\top\right)^{-1} \\ &= G(z) - \frac{\theta\, G(z)\, uv^\top\, G(z)}{1 + \theta\, v^\top G(z) u}. \end{aligned} \quad (3)$$

The poles of $G_{\mathbf{M}}(z)$ correspond to the eigenvalues of $\mathbf{M}$. Thus, the condition:

$$1 + \theta\, v^\top G(z) u = 0,$$

determines the emergence of new eigenvalues due to the low-rank perturbation.

### *B. The BBP Transition*

The BBP transition, as described in [1], characterizes the phenomenon by which a low-rank perturbation gives rise to an eigenvalue that detaches from the bulk spectrum of a random matrix.

Assume that for large $N$ the quadratic form $v^\top G(z) u$ concentrates around the Stieltjes transform $m(z)$. Then the condition for the emergence of an outlier eigenvalue is:

$$1 + \theta\, m(\lambda_{\text{out}}) = 0,$$

or equivalently,

$$\theta\, m(\lambda_{\text{out}}) = -1.$$

For $z > \lambda_+$ (the upper edge of the MP bulk), the Stieltjes transform is typically given by:

$$m(z) = \frac{z - \sqrt{z^2 - 4\sigma^2}}{2\sigma^2}.$$

Solving the equation:

$$\theta\, \frac{z - \sqrt{z^2 - 4\sigma^2}}{2\sigma^2} = -1,$$

yields an expression for the outlier eigenvalue. After rearrangement and under appropriate conditions (typically $\theta > \sigma$), the solution takes the form:

$$\lambda_{\text{out}} = \theta + \frac{\sigma^2}{\theta}.$$

This analysis shows that when the strength $\theta$ of the perturbation exceeds a critical threshold, an eigenvalue (or singular value, in the non-symmetric case) detaches from the bulk. This phenomenon is the hallmark of the BBP transition and underpins the emergence of *intruder dimensions* in LoRA.

### *C. Extension to Rank-$r$ Updates*

For a general low-rank update:

$$\Delta \mathbf{W}_{\text{LoRA}} = \sum_{k=1}^{r} \gamma_k \, p_k q_k^\top,$$

each singular value $\gamma_k$ (with corresponding singular vectors $p_k$ and $q_k$) behaves analogously to the rank-1 case. Under the assumption that the spikes $\gamma_k$ are sufficiently separated and exceed the threshold determined by the upper edge $\sqrt{\lambda_+}$ of the MP distribution, the condition:

$$1 + \gamma_k \, m(\lambda_{\text{out},k}) = 0$$

yields outlier singular values:

$$\lambda_{\text{out},k} \approx \gamma_k + \frac{\sigma^2}{\gamma_k}, \quad \text{for } \gamma_k > \sqrt{\lambda_+}.$$

If the singular vectors $p_k$ and $q_k$ associated with these spikes are nearly orthogonal to the pre-trained singular vectors of $\mathbf{W}_{\text{pre}}$, then these outlier singular values represent new, *intruder dimensions* that disrupt the alignment of the fine-tuned model with its pre-trained spectral basis.

## V. Mitigation Strategies

Given the potential negative impact of intruder dimensions on model generalization and stability, it is important to explore strategies that mitigate their formation.

### *A. Increasing the Rank of the Update*

One intuitive approach is to increase the rank $r$ of the LoRA update. By distributing the update energy over a higher-dimensional subspace, the influence of any single spike is diminished. However, it has been observed that low-rank updates often suffer from a *small effective rank*—that is, the actual expressive capacity of the update is much lower than its nominal rank. This phenomenon limits the model's adaptability, as important directions in the parameter space may remain underrepresented. Although increasing $r$ can partially alleviate the formation of dominant outliers and mitigate the issue of a small effective rank, it comes at the cost of increased parameter count and computational complexity.

### *B. Rank-Stabilized LoRA*

A major drawback of standard LoRA updates is that their singular vectors may not align well with the pre-trained weight matrix, leading to spectral distortions and the emergence of intruder dimensions. Additionally, low-rank updates often suffer from a **small effective rank**, meaning that much of the update energy is concentrated in a few dominant directions rather than being evenly distributed.

**Rank-stabilized LoRA** addresses these issues by enforcing constraints on the update matrices $\mathbf{B}$ and $\mathbf{A}$, ensuring that the update energy is spread more evenly across multiple directions. A common approach is to impose orthogonality conditions:

$$\mathbf{B}^\top \mathbf{B} = \mathbf{I}_r, \quad \mathbf{A}^\top \mathbf{A} = \mathbf{I}_r,$$

which prevents singular values from collapsing into a few dominant spikes. This not only reduces spectral outliers but also increases the effective rank of the update, ensuring that more directions in parameter space are utilized for adaptation.

$$\Delta \mathbf{W}_{\mathbf{LoRA}} = \frac{\alpha}{r} \mathbf{B} \mathbf{A}^\top$$

By stabilizing the update rank, this method helps prevent the formation of large singular value spikes, mitigates small effective rank issues, and improves the robustness of LoRA, making it behave more like full fine-tuning in terms of spectral properties.

### *C. Rank Stabilization via Spectral Fine-Tuning*

An alternative and promising approach is to perform spectral fine-tuning. This method leverages the close connection between matrix rank and spectral representation. Specifically, consider the Singular Value Decomposition (SVD) of the pre-trained weight matrix:

$$\mathbf{W} = \mathbf{U}\mathbf{S}\mathbf{V}^\top.$$

[9] defines two mechanisms for spectral fine-tuning:

1) **Additive Spectral Adapter:**

$$\text{Adapter}_A(\mathbf{W}) := [\mathbf{U}_1 + \mathbf{A}_U \, \mathbf{U}_2] \, \mathbf{S} \, [\mathbf{V}_1 + \mathbf{A}_V \, \mathbf{V}_2]^\top.$$

2) **Rotational Spectral Adapter:**

$$\text{Adapter}_R(\mathbf{W}) := [\mathbf{U}_1 \, \mathbf{R}_U \, \mathbf{U}_2] \, \mathbf{S} \, [\mathbf{V}_1 \, \mathbf{R}_V \, \mathbf{V}_2]^\top.$$

In both cases, the approach is to fine-tune the dominant singular components of $\mathbf{W}$ rather than perturbing the entire matrix arbitrarily. By doing so, the fine-tuning process remains aligned with the spectral structure inherited from pre-training, thereby suppressing the formation of harmful intruder dimensions.

### *D. Discussion of Mitigation Approaches*

The additive and rotational spectral adapters offer complementary ways to control the spectral update. The additive method perturbs the singular vectors directly by adding a correction term, while the rotational method applies a transformation that rotates the singular vectors within the appropriate subspace. Both methods are designed to keep the

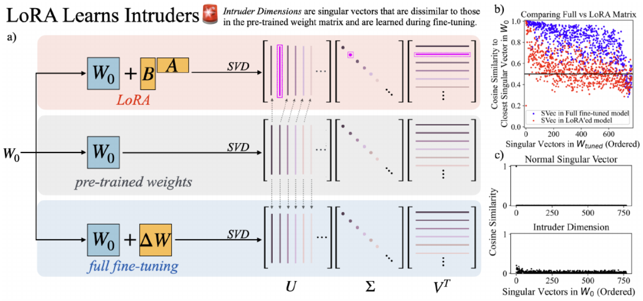
LoRA Learns Intruders
Intruder Dimensions are singular vectors that are dissimilar to those in the pre-trained weight matrix and are learned during fine-tuning.
a)
$W_0$
$W_0$ + $B$ $A$
LoRA
SVD
$W_0$
pre-trained weights
SVD
$W_0$ + $\Delta W$
full fine-tuning
SVD
$U$
$\Sigma$
$V^T$
b)
Comparing Full vs LoRA Matrix
Cosine Similarity to Closest Singular Vector in $W_0$
1.0
0.8
0.6
0.4
0.2
0.0
0
200
400
600
Singular Vectors in $W_{tuned}$ (Ordered)
SVec in Full fine-tuned model
SVec in LoRA'ed model
c)
Normal Singular Vector
Intruder Dimension
Cosine Similarity
1
0
0
250
500
750
Singular Vectors in $W_0$ (Ordered)

Fig. 2. Characterizing structural differences between solutions learnt by LoRA Vs full Fine- tuning. a) [4] measure the changes to the SVD of the pre-trained weights made during fine-tuning. They observe intruder dimensions introduced by LoRA in top ranking singular vectors but by full fine- tuning. b) Comparing a matrix fine-tuned with full fine-tuning or LoRA. c) Comparing a normal singular vs an intruder dimension to all pre-trained singular vectors.

update closely aligned with the pre-trained structure, and our experimental observations suggest that they can significantly reduce the presence of outlier singular values in the fine-tuned model.

### E. Discussion

The experimental results in [4] corroborate our RMT-based theoretical analysis. Full fine-tuning preserves the spectral structure of the pre-trained matrix, while LoRA, due to its low-rank nature, introduces intruder dimensions via the BBP transition. Mitigation strategies that either increase the effective rank or apply spectral alignment successfully reduce the formation of these spectral outliers, thereby improving the generalization performance and stability of the fine-tuned model.

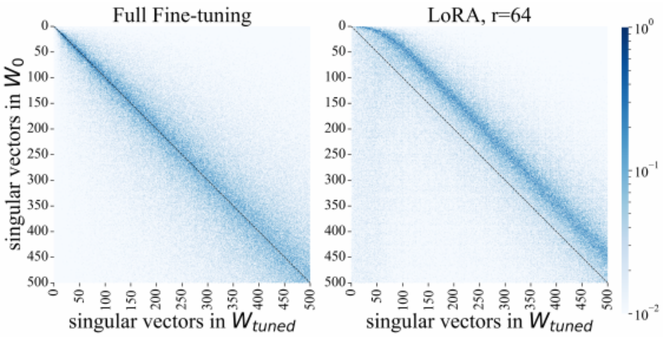

Fig. 3. Spectral dissimilarities between full fine- tuning and LoRA. Similarity matrix of pre- and post-fine-tuning singular vectors of the weight matrices to characterize spectral differences introduced upon fine-tuning, in a representative example for LLaMA-2 fine-tuned on Magicoder. Full fine-tuning retains most of the pre-training structure; the diago- nal shift in LoRA corresponds to the introduction of intruder dimensions, color shows cosine similiarity

## VI. Conclusion and Future Work

In this paper, we have presented a comprehensive analysis of the spectral properties of LoRA and full fine-tuning using Random Matrix Theory. Our investigation reveals that: noitemsep

- LoRA, by virtue of its low-rank update, introduces intruder dimensions through the BBP phase transition. These intruder dimensions represent new singular vectors that are misaligned with the pre-trained spectral basis.
- Full fine-tuning, in contrast, employs a dense, small perturbation that preserves the original spectral structure (the MP bulk).
- Mitigation strategies such as increasing the rank of the LoRA update and applying spectral fine-tuning (via additive or rotational adapters) can significantly reduce the formation of spectral outliers, thereby aligning LoRA's behavior more closely with that of full fine-tuning.

Our work provides a rigorous theoretical foundation for understanding the spectral differences between parameter-efficient fine-tuning methods and traditional full fine-tuning. These insights have important implications for the design of robust adaptation techniques in deep learning, especially in the context of sequential learning and transfer learning tasks.

### A. Future Work

Future research directions include: noitemsep

- Extending the analysis to non-asymptotic settings and incorporating higher-order perturbation effects.
- Investigating the impact of different distributions for the pre-trained weight matrix.
- Exploring adaptive mechanisms for rank stabilization that can dynamically adjust to the spectral properties of the model during fine-tuning.
- Empirical studies on large-scale models in practical applications to validate the theoretical predictions further.

## Acknowledgements

We would like to thank our colleagues and the High-Dimensional Probability group for valuable discussions.

## References

[1] Baik, J., Ben Arous, G., & Péché, S. (2005). *Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices.*

[2] Johnstone, I. M. (2001). *On the distribution of the largest eigenvalue in principal components analysis.*

[3] Noiry, N. (2020). *Spectral Measures of Spiked Random Matrices.*

[4] Shuttleworth, R., Andreas, J., Torralba, A., & Sharma, P. (2024). *LoRA vs Full Fine-tuning: An Illusion of Equivalence.*

[5] Couillet, R. & Liao, Z. (2022). *Random Matrix Methods for Machine Learning.* Cambridge University Press.

[6] Livan, G., Novaes, M., & Vivo, P. (2017). *Introduction to Random Matrices - Theory and Practice.* ArXiv. https://doi.org/10.1007/978-3-319-70885-0

[7] Nadakuditi, R. R. (2011). The singular values and vectors of low rank perturbations of large rectangular random matrices. ArXiv. https://arxiv.org/abs/1103.2221

[8] Hu, E. J., Shen, Y., Wallis, P., Li, Y., Wang, S., Wang, L., & Chen, W. (2021). *LoRA: Low-Rank Adaptation of Large Language Models.* ArXiv. https://arxiv.org/abs/2106.09685

[9] Zhang, F., & Pilanci, M. (2024). *Spectral Adapter: Fine-Tuning in Spectral Space.* ArXiv. https://arxiv.org/abs/2405.13952