

# Table of Content

<b>1. Preliminaries</b>	<b>1</b>
1.1. Introduction . . . . .	1
1.2. Neural Networks (NNs) . . . . .	2
1.2.1. Feedforward Neural Networks . . . . .	2
1.2.2. Other Architectures for Neural Networks . . . . .	3
1.2.2.1. Convolutional NNs (CNNs) . . . . .	3
1.2.2.2. Recurrent Neural Networks . . . . .	4
1.3. Training Neural Networks and Machine Learning Models . . . . .	4
1.3.1. Gradient Descent . . . . .	5
1.3.2. Stochastic Gradient Descent and Variations . . . . .	6
1.4. Reproducing Kernel Hilbert Spaces . . . . .	7
1.4.1. Building a RKHS . . . . .	8
1.4.2. Interpolation and Adjustment . . . . .	8
1.5. Functional Gradient Descent . . . . .	10
<b>2. Different Parametrizations of Shallow Neural Networks</b>	<b>11</b>
2.1. Mean Field Regime . . . . .	12
2.2. Neural Tangent Kernel Regime . . . . .	13
2.2.1. Lazy Training in Neural Networks . . . . .	15
<b>3. Main Results</b>	<b>16</b>
3.1. The limiting dynamics of the empirical measure . . . . .	16
3.2. The Limit of the Neural Network . . . . .	19
<b>4. The NTK Regime through the lens of mean field models</b>	<b>22</b>
4.1. Training Dynamics . . . . .	24
4.1.1. Technical Lemmas . . . . .	25
4.2. Existence of Solutions for the SDE . . . . .	31
4.3. Tightness of the laws of the empirical measure process . . . . .	33
4.3.1. First Part of the Proof . . . . .	33
4.3.2. Second Part of the Proof . . . . .	46
4.4. The PDE limit . . . . .	53
4.4.1. Identification of the Limit . . . . .	53
4.4.2. Convergence to the Limit . . . . .	57
4.5. Uniqueness of solutions for the PDE limit . . . . .	67
4.6. The solution with Xavier Initialization and a Construction of the NTK . . . . .	77
<b>5. The Neural Network with Xavier Initialization and SGD training</b>	<b>80</b>

5.1. Identifying the limit . . . . .	81
5.2. Studying the Limiting Dynamic with Xavier initialization . . . . .	90
<b>6. Conclusion</b>	<b>95</b>
<b>Bibliography</b>	<b>96</b>
<b>1. Annex</b>	<b>99</b>
1.1. Controlling the moments of the parameters . . . . .	99