

Contents

1	Introduction	1
1.1	Related Work	3
1.1.1	Spiking Neural Network Application and Formulation	3
1.1.2	Optimisation Criteria Based on Information Theory	4
1.2	Research Problem	4
1.3	Hypothesis	5
1.3.1	Parameter Model Definition on Spiking Models	5
1.3.2	Fair Comparison	6
1.4	Research Questions	6
1.5	Objectives	6
1.5.1	General Objective	6
1.5.2	Specific Objectives	6
1.6	Methodology	6
1.7	Contributions	7
1.8	Thesis Outline	7
2	Theoretical Framework	8
2.1	Recurrent Neural Networks (RNNs)	9
2.1.1	Vanilla RNN (V-RNN)	9
2.1.2	Long Short-Term Memory (LSTM)	10
2.1.3	Gated Recurrent Unit (GRU)	11
2.1.4	RNN Architectures	12
2.2	Spiking Recurrent Neural Networks (SRNNs)	13
2.2.1	Vanilla SRNN (V-SRNN)	14
2.2.2	Cramer SRNN (C-SRNN)	15
2.2.3	Adaptive SRNN (A-SRNN)	16
2.3	Supervised Learning	17
2.3.1	Cross-Entropy Optimisation Criterion	17
2.3.2	Evaluation Metric: Accuracy	17
2.4	Training RNNs	18
2.4.1	Vanishing and Exploding Gradient Problem	19
2.4.2	Gradient Clipping	19
2.4.3	Reducing the Learning Rate at the Plateau	20
2.5	Information Theory	20
2.5.1	Random Variables	20

2.5.2	Entropy	21
2.5.3	Mutual Information	21
2.5.4	Information Bottleneck Principle	21
3	Methodology	23
3.1	Experimental Data	23
3.1.1	Parity Bit Problem or Sequential XOR	23
3.1.2	Spiking Heidelberg Digits (SHD)	24
3.2	Experimental Framework	25
3.2.1	Recurrent Neural Networks Graph Representation	25
3.2.2	Supervised Learning setup for Multilayer SRNNs and traditional RNNs	28
3.2.3	Backpropagation in SRNNs	31
3.2.4	Training Setup	32
3.2.5	Evaluation Procedure	33
3.2.6	Final experimental flow	33
3.3	SRNN Model Parameters	35
3.3.1	Analysis of Parameters	35
3.3.2	Analysis of Surrogate Gradient	36
3.3.3	Initialisation of Learnable Parameters (Weights)	36
3.3.4	Proposed method to set SRNNs hyperparameters	36
3.4	Memory in SRNNs	39
3.4.1	Categorisation of SRNNs based on their type of memory	39
3.4.2	Adaptation of Two-Level External Memory in SRNNs	40
3.5	Information Bottleneck in SRNNs	42
3.5.1	Proposed Mutual Information Estimator Based on the Central Limit Theorem	43
4	Baseline Results	44
4.1	Parity Bit Problem	44
4.1.1	Discussion	47
4.2	SHD Problem Baselines	47
4.2.1	Discussion	49
5	SRNN Model Hyperparameters	51
5.1	Heuristic Method Exploration	51
5.1.1	Discussion	54
5.2	SRNNs vs Traditional RNNs	56
5.2.1	Discussion	60
5.3	Literature fixed Parameters vs Heuristic Method	62
5.3.1	Discussion	63
6	Memory and Optimisation Criterion	65
6.1	Two-Level External Memory	65
6.1.1	Discussion	67
6.2	Space Information Bottleneck on V-SRNN	69
6.2.1	Discussion	72
7	Conclusions	73

7.1 Final Comments and Future Work	75
Bibliography	75
Appendices	83
A Linear Transformation Approximates a Normal Distribution	84
B Input Current PDF Approximation	85
C Maximum Scores of Baselines	86