

UNITARY EVOLUTION RECURRENT NEURAL NETWORKS

Martin Arjovsky *

Universidad de Buenos Aires
{marjovsky}@dc.uba.ar

Amar Shah *

Cambridge University
{as793}@cam.ac.uk

Yoshua Bengio

Universite de Montréal, CIFAR Senior Fellow
{yoshua.bengio}@gmail.com

ABSTRACT

Recurrent neural networks (RNNs) are notoriously difficult to train. When the eigenvalues of the hidden to hidden weight matrix deviate from absolute value 1, optimization becomes difficult due to the well studied issue of vanishing and exploding gradients, especially when trying to learn long-term dependencies. To circumvent this problem, we propose a new architecture that learns a unitary weight matrix, with eigenvalues of absolute value exactly 1. We construct an expressive unitary weight matrix by composing several structured matrices that act as building blocks with parameters to be learned. Optimization of this parameterization becomes feasible only when considering hidden states in the complex domain. We demonstrate the potential of this architecture by achieving state of the art in several hard tasks involving very long-term dependencies.

*Indicates first authors. Ordering determined by coin flip.