

IMEXnet - A Forward Stable Deep Neural Network

Eldad Haber, Keegan Lensink, Eran Treister and Lars Ruthotto

Jun 2019

Outline

- ▶ Why Implicit
- ▶ Implicit Explicit
- ▶ Some results

Why Implicit

- ▶ For CNN's - depth is connected to field of view
- ▶ Stability of the standard networks can be limited
- ▶ Vanishing/Exploding gradients

Goal: Develop a method that can deal with those problems

Deep Networks and ODE's

$$\dot{\mathbf{Y}} = \sigma(\mathbf{KY} + \mathbf{b}) \quad \leftrightarrow \quad \mathbf{Y}_{j+1} = \mathbf{Y}_j + h\sigma(\mathbf{K}_j\mathbf{Y}_j + \mathbf{b}_j).$$

- ▶ Deep Residual Networks equivalent to Forward Euler for ODE's
- ▶ Forward Euler have limitation on stability
- ▶ Require many steps to converge

Semi-Implicit methods

Different stable integration technique that allows large steps

$$\dot{\mathbf{Y}} = \sigma(\mathbf{K}\mathbf{Y} + \mathbf{b}) \quad \leftrightarrow \quad \mathbf{Y}_{j+1} = (\mathbf{I} + h\mathbf{K}_j)^{-1} (\mathbf{Y}_j + h\sigma(\mathbf{K}_j\mathbf{Y}_j + \mathbf{b}_j) - \mathbf{K}_j\mathbf{Y}_j).$$

Implicit methods are used for

- ▶ Computational Fluid Dynamics
- ▶ Computational Electromagnetics
- ▶ Nonlinear dynamics
- ▶ Computer graphics

Semi-Implicit methods

Come to our poster and see how we apply these networks to many data sets