جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

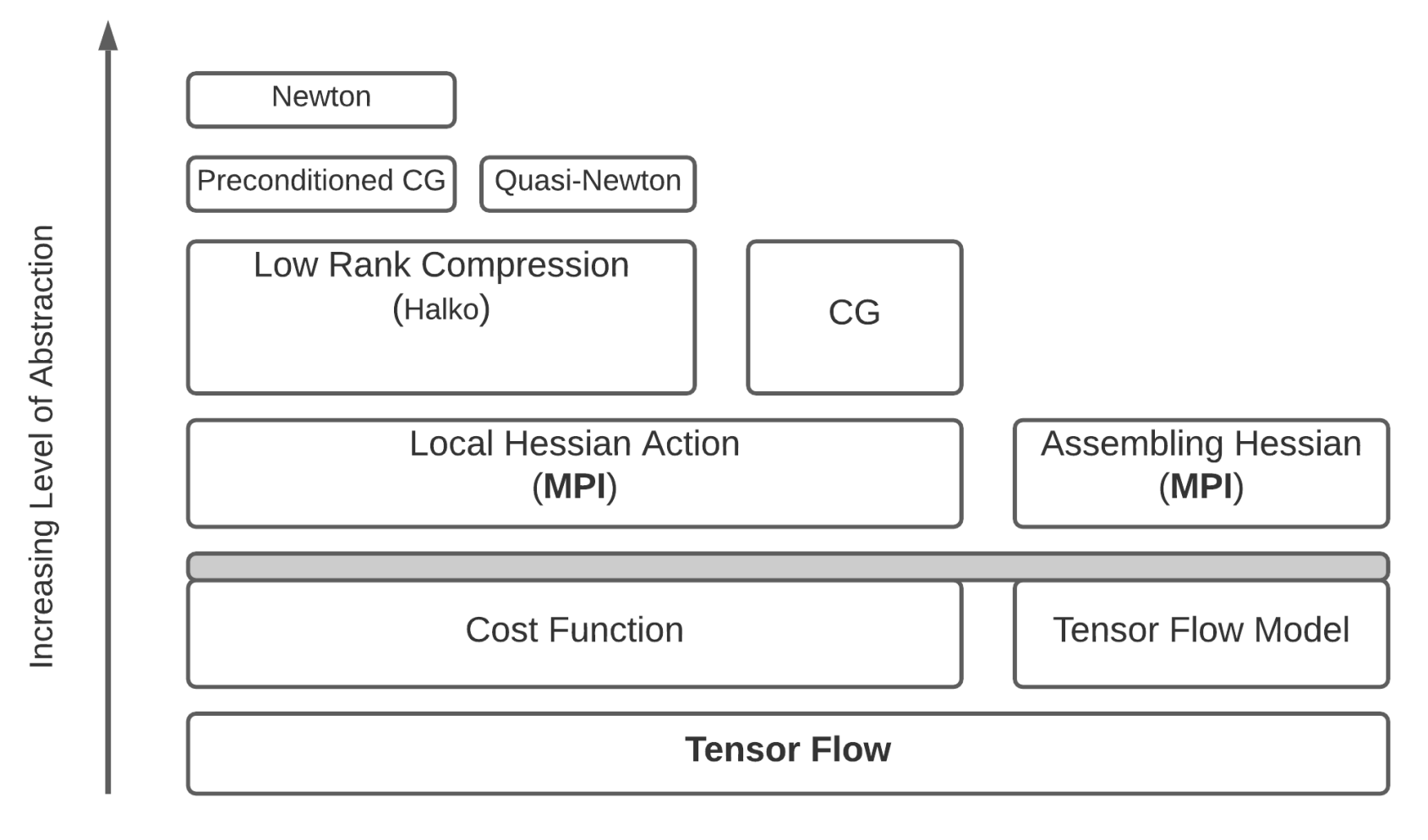
Hessian Compression in Machine Learning

Tensorflow, MPI, Random Sampling method

November 2, 2021

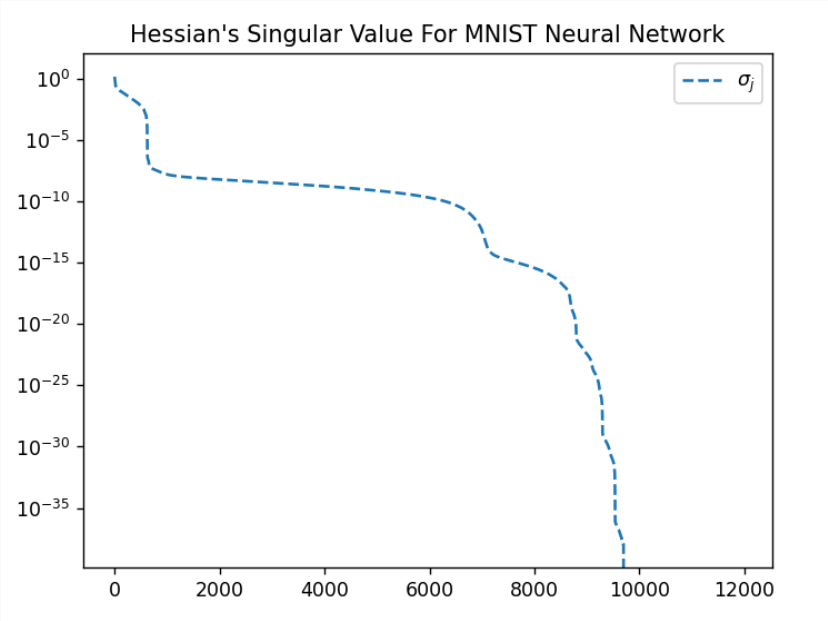
Developed Components

Developed Components



Hessian Compression

Algorithms presented in Finding Structure with Randomness: Probabilistic Algorithms for Constructing Approximate Matrix Decompositions by Halko, N. and Martinsson, P. G. were implemented to compress the Hessian.



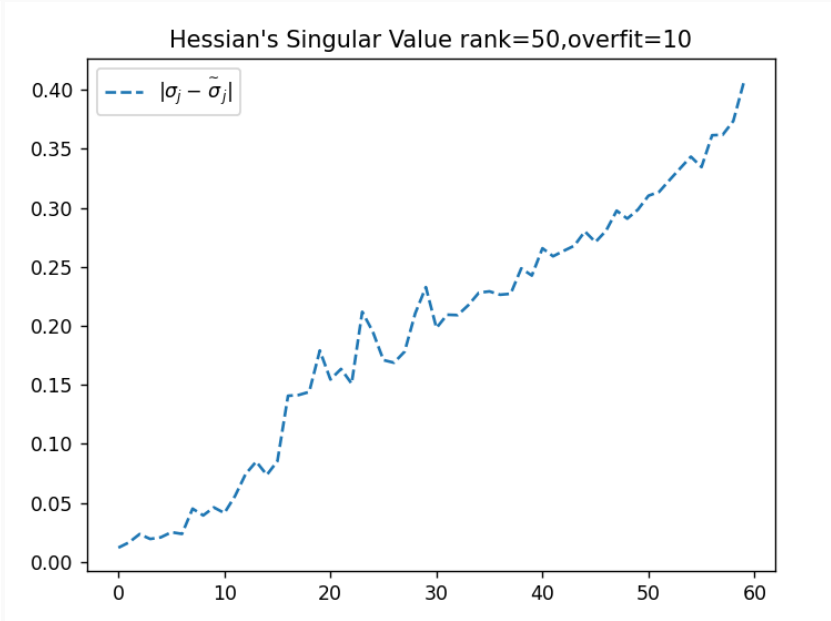


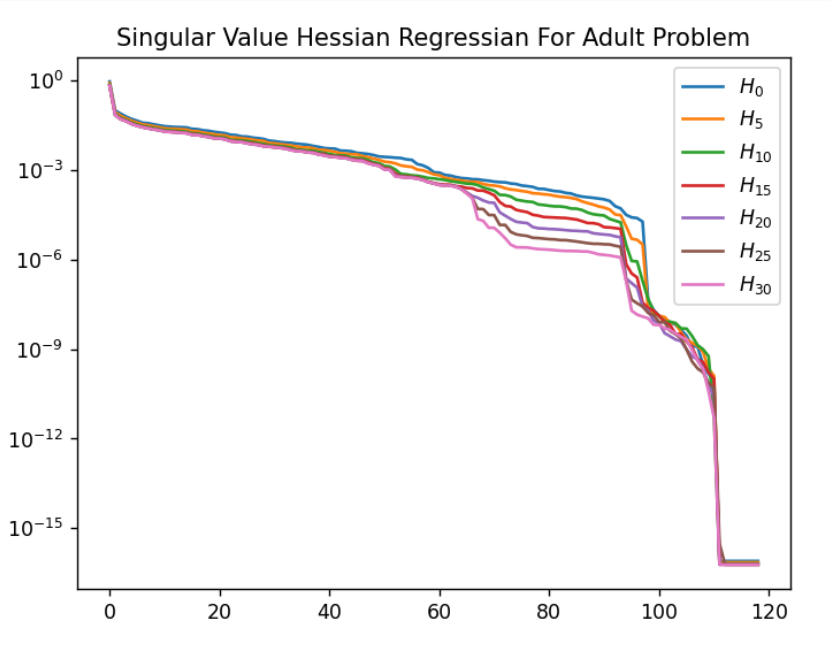
Figure 1: The graph refers to the Hessian of Neural Network trained on MNIST data set with a 15 node hidden layer with sigmoid activation function.

Convexity – Adult Salary Problem

When working with convex energy functional one can use the components developed in order to minimize the energy of the problem.

In particular when working with convex energy second order method, such as the Newton method will converge to global minimums. In particular as a test problem we considered the adult slary problem with the following energy,

$$E(\vec{x}) = \frac{1}{m}\sum_{i=1}^{m} \log\left(1 + \exp\left(-b_j \vec{a_j}^T \vec{x}\right)\right)$$
$$\forall\, x \in \mathbb{R}^d$$



where d is the feature number and $\vec{a_j}$ are the data while b_j are the labels.

In particular the following Newton method was implemented,

$$\vec{x}_{n+1} = \vec{x}_n - \gamma Hf(\vec{x}_n)^{-1} \nabla f(\vec{x}_n).$$

Future Work

Future Question

- Meaning full energy minimization problem,
  - Convex Energy,
  - High Dimension,
  - Real life application.
- Training Jinchao Xu NN-FEM using second order method to minimize the Dirichlet energy functional.
- Develop possible connections between Least Square Finite Elements and NN-FEM.

Thank you !