Department of Informatics

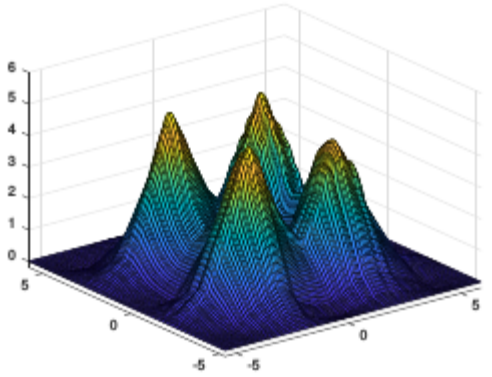
Technical University of Munich

Poster C-03

HPC-Suitable Data Structures for Machine Learning and Other Applications of Adaptive Sparse Grids

Paul - Cristian Sârbu
paul.c.sarbu@tum.de

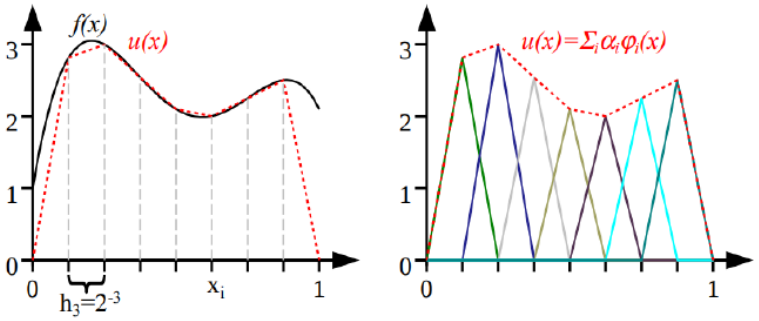
Supervisor: Univ.-Prof. Dr. Hans-Joachim Bungartz



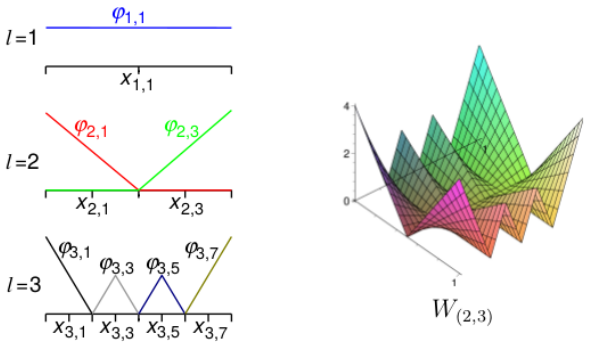
International HPC Summer School 2018
July 8-13, Ostrava, Czech Republic

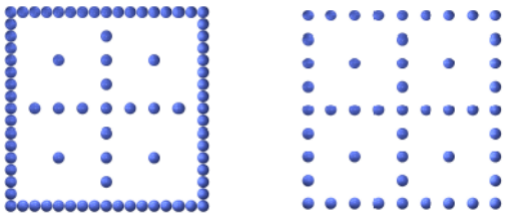
Sparse Grids (SG) in a Nutshell

Efficient representation, interpolation and quadrature of high-dimensional functions

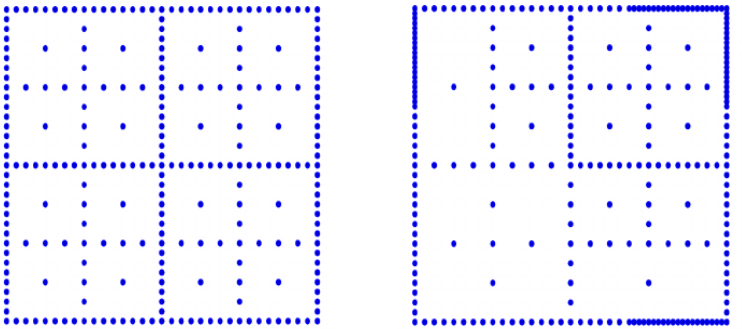


Different basis functions give different grids





Adaptive refinement used for high accuracy by fitting to the underlying data



General Sparse Grids Toolbox developed at the Chair of Scientific Computing, Munich (C++ with wrappers for Python, Java, Matlab)

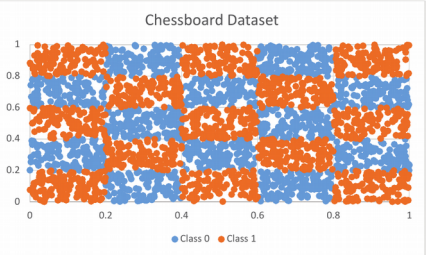
Applications:

- function interpolation
- quadrature
- partial differential equations
- data mining & machine learning
- uncertainty quantification
- function optimization
- ...

Results & Work in Progress (WiP)

- Optimization of an adaptive sparse grid-based regression / classification kernel for many-core architectures (KNL) - **Done**

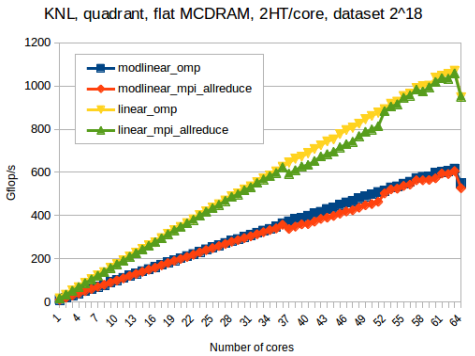
  Project: Intel Parallel Computing Center at Leibniz Supercomputing Center (LRZ)

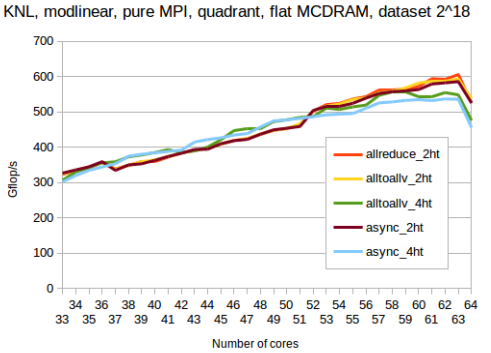


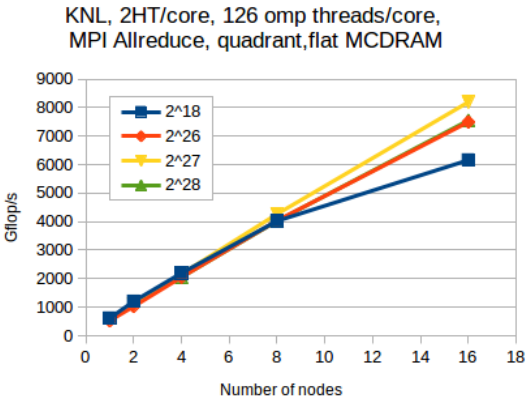
Datasets of up to 2^{28} instances (20GB)

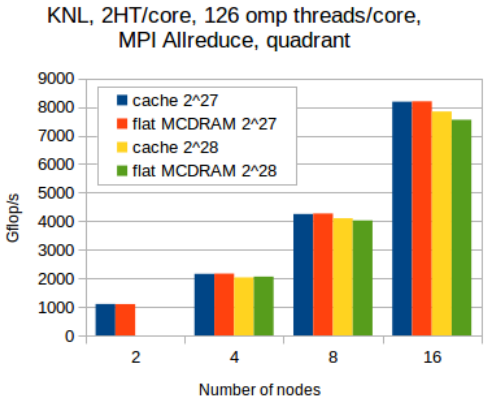
Speedups: 1.6-1.7x vs. Haswell

Next: Skylake & much larger simulations

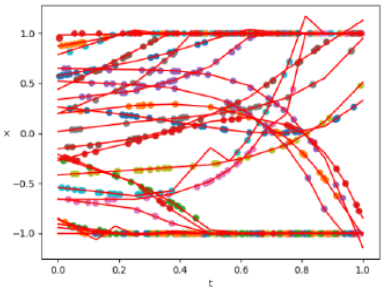


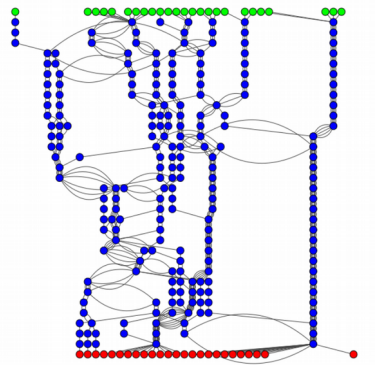






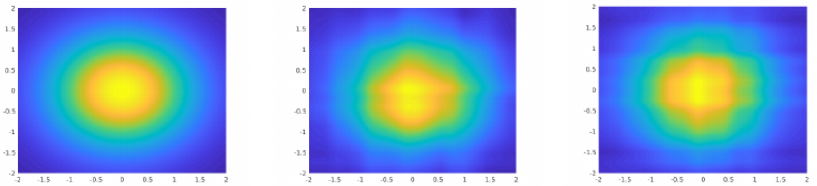
- Applications of sparse grid-based interpolation and regression - **WiP**
  - Multidimensional SG interpolation for trajectory clustering
  - Time series prediction using regression (on a method by Bohn, Gribel et al.)





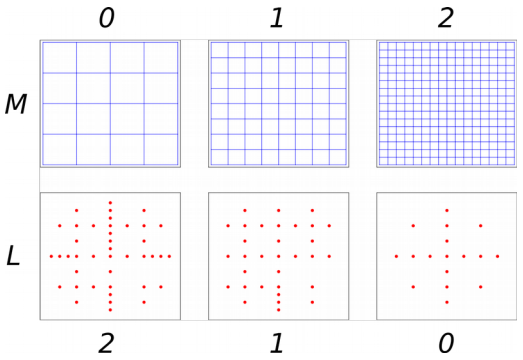
Results & Work in Progress (WiP) – *cont.*

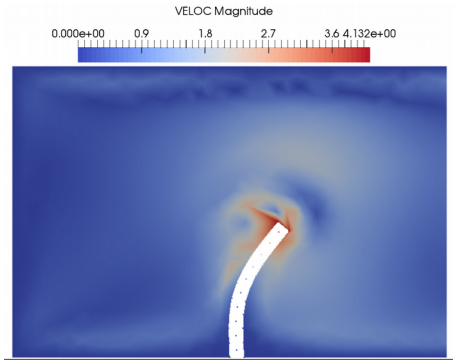
- Applications of sparse grid density estimation (SGDE) - **WiP**
  - SGDE techniques for clustering-based stochastic collocation:
    - → adapt method from literature to parallel SG treatment
    - → apply several approaches to use the cheaply computed SGDE for clustering
    - → handle uncertainty in the data (e.g. representative clustering)

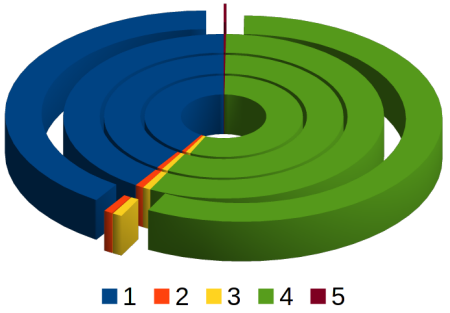


- Other applications of efficient adaptive sparse grids – **more to come**
  - Uncertainty quantification for fluid structure interaction problems:

Farcaş IG., **Sârbu P.C.**, Bungartz HJ., Neckel T., Uekermann B. (2018) **Multilevel Adaptive Stochastic Collocation with Dimensionality Reduction**. In: Garcke J., Pflüger D., Webster C., Zhang G. (eds) Sparse Grids and Applications - Miami 2016. Lecture Notes in Computational Science and Engineering, vol 123. Springer, Cham

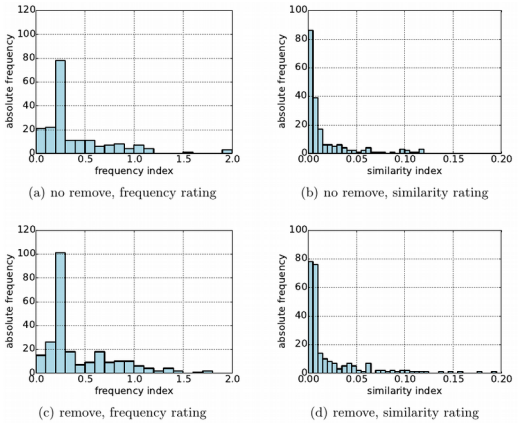


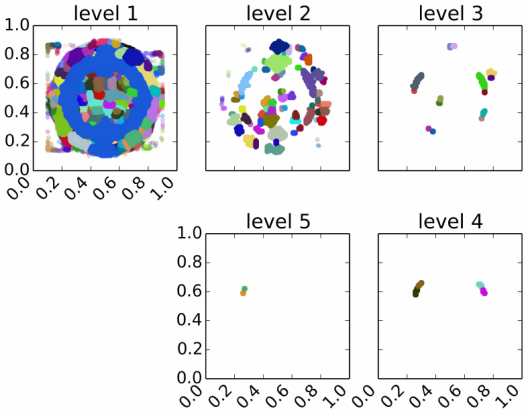


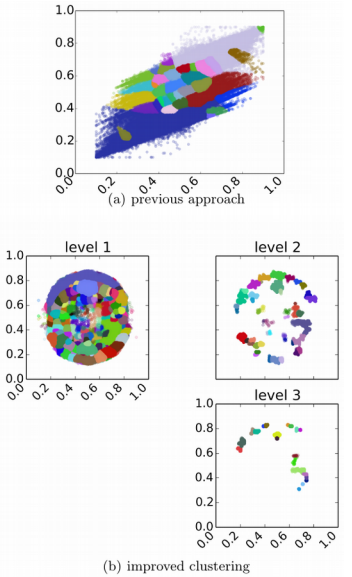


Future work

- SGDE in UQ problems with unknown or highly coupled distributions
- Large scale runs with the optimized regression/classification kernel for different applications
- Scalable batch learner for SGDE with MPI parallelization
- Scalable recommender system using SGDE clustering
  - ➔ renewing legacy code
  - ➔ Implementing new and efficient parallel algorithms







- Expanding SG++: new clustering module, new HPC python module for rapid prototyping