Sparse Grid Methods for Uncertainty Quantification

Michael Griebel

University of Bonn and Fraunhofer SCAI

Joint work with Alexander Rüttgers

1. Sparse grids
   - Construction principles and properties
   - Optimal sparse grids
   - Adaptive combination method
2. Application
   - Multi-scale viscoelastic flows

Motivation

- Numerical methods in uncertainty quantification:
  - Galerkin approach
  - Collocation technique
  - Discrete projection
- Needed on stochastic/parameter domain:
  - Approximation of integrals
  - Interpolation, especially for collocation
- Simple domains with product structure: $[-a,a]^d$, $\mathbb{R}^d$
- Issue: high- or even infinite-dimensional problems

Curse of dimension

- $f:\Omega^{(d)} \to \mathbb{R}, \quad f \in V^{(r)},$ r isotropic smoothness
- Bellmann ´61: curse of dimension $\quad M = \#\text{dof}$

$$\| f - f_M \|_{H^s} = C(d) \cdot M^{-r/d} \, | f |_{H^{s+r}} = O(M^{-r/d})$$

- Find situations where curse can be broken ?
- Trivial: restrict to $r = O(d)$

$$\| f - f_M \| = O(M^{-cd/d}) = O(M^{-c})$$

  but practically not very relevant
- In any case: some smoothness changes with d
  or importance of coordinates decays successively
  (e.g. after suitable nonlinear transformation)

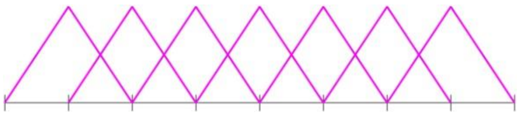
Sparse grid approach

- Basic principles:
  - 1-dim multilevel series expansion with proper decay
  - d-dim product construction
  - Trunctation of resulting multivariate expansion

- Effect:
  - reduction of cost complexity
  - nearly same accuracy as „full“ product
  - necessary: certain smoothness requirements
  - adaptivity for detection of lower-dimensional manifolds

Simple example: Hierarchical basis

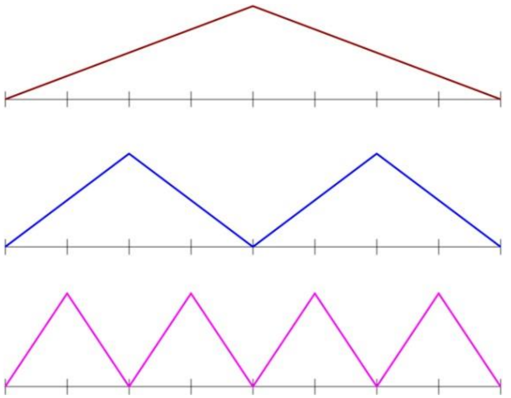
V_3



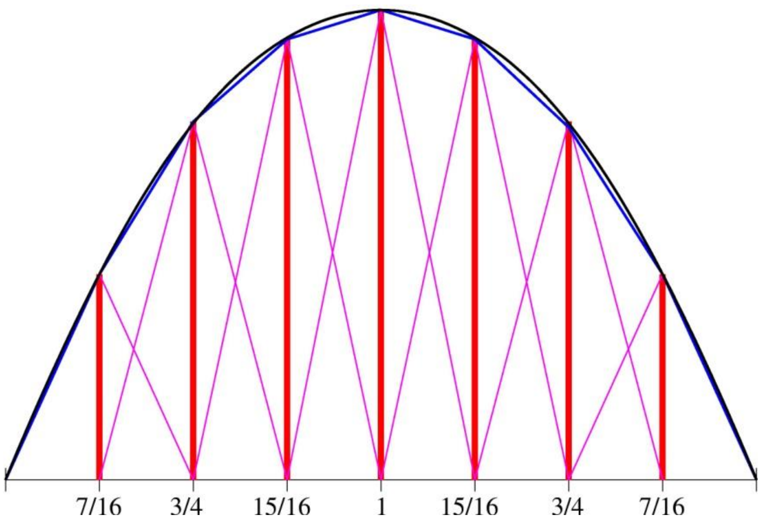
$l=1$ W_1

$l=2$ W_2

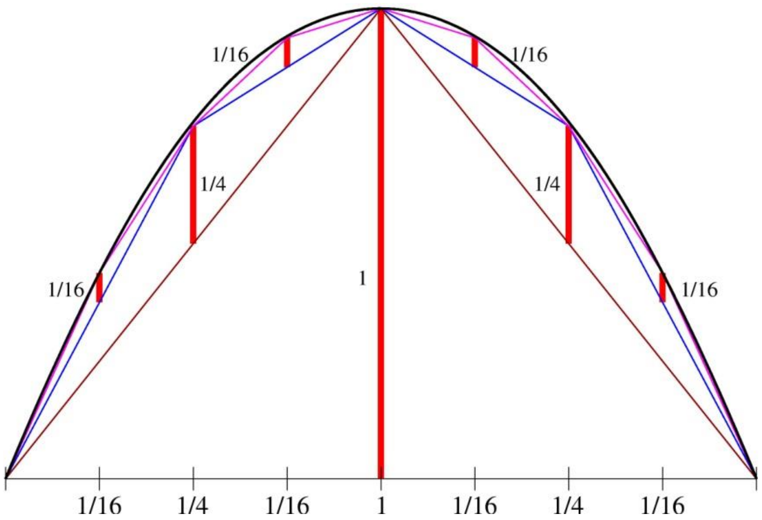
$l=3$ W_3



parabola $f(x) = -(x-1)(x+1)$ in [-1,1]



conventional coefficients
no decay from level to level



hierarchical coefficients
decay by ¼ from level to level

Tensor product hierarchical basis

Generalization to higher dimension by tensor product

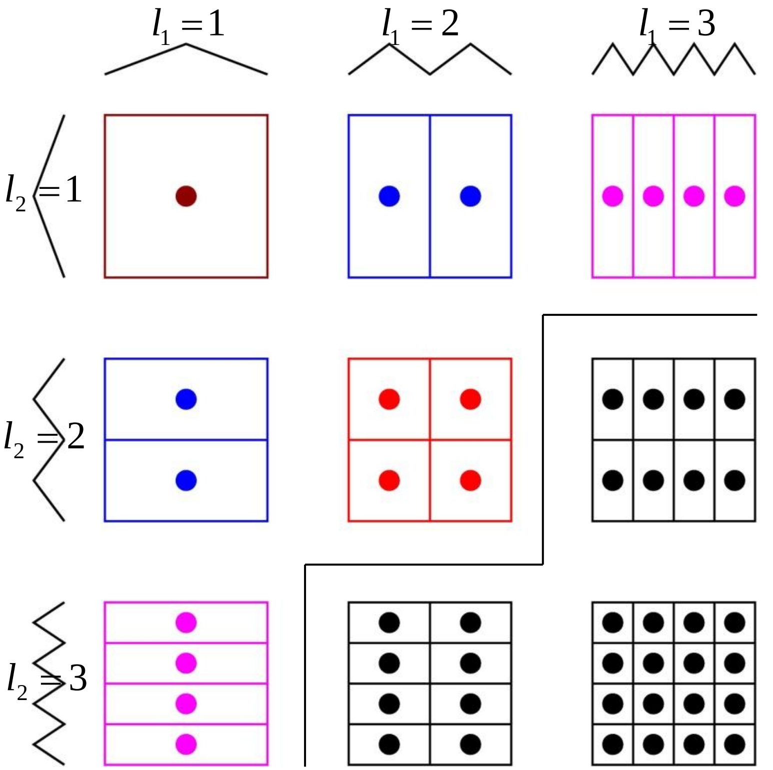


Table of subspaces $W_{l_1 l_2}$

$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$	$\frac{1}{16}$	$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$
$\frac{1}{64}$	$\frac{1}{16}$	$\frac{1}{64}$	$\frac{1}{4}$	$\frac{1}{64}$	$\frac{1}{16}$	$\frac{1}{64}$
$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$	$\frac{1}{16}$	$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$
$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{16}$	1	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{16}$
$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$	$\frac{1}{16}$	$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$
$\frac{1}{64}$	$\frac{1}{16}$	$\frac{1}{64}$	$\frac{1}{4}$	$\frac{1}{64}$	$\frac{1}{16}$	$\frac{1}{64}$
$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$	$\frac{1}{16}$	$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$

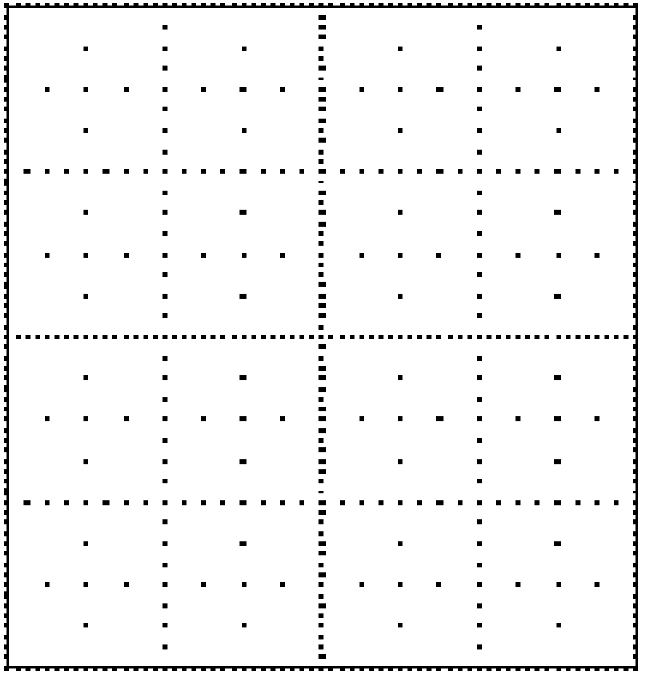
decay in x- and y-direction by 1/4

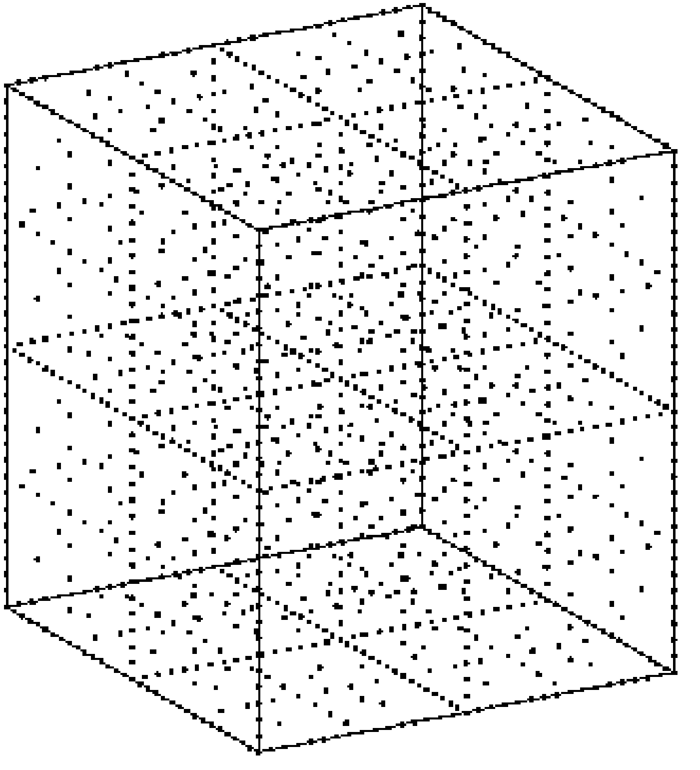
decay in diagonal direction by 1/16

Idea:

Omit points with small associated hierarchial coefficient values

Regular sparse grids





Properties of regular sparse grids

$N \cong 2^n$	Sparse grids		Full grids	
Cost:	$O(N \log(N)^{d-1})$	instead of	$O(N^d)$	
Accuracy:	$O(N^{-2} \log(N)^{d-1})$		$O(N^{-2})$	L_2-norm
Smoothness:	$\left\| \frac{\partial^{2d} f}{\partial x_1^2 \ldots \partial x_d^2} \right\| \leq c$		$\left\| \sum_{i=1}^{d} \frac{\partial^2 f}{\partial x_i^2} \right\| \leq c$	
Space, seminorm:	$H^2_{mix}, \| f \|_{2,mix}$		$H^2, \| f \|_2$	

Mitigates the curse of dimension of conventional full grids
Note: Higher regularity in mixed derivative, ~d

For wavelets, general stable multiscale systems: $O(N^{-2} (\log N)^{(d-1)/2})$

History of regular sparse grids

Re-invented several times:

1957	Korobov, Babenko	hyperbolic cross points
1963	Smolyak	
1970	Gordon	blending method
1980	Delvos, Posdorf	Boolean interpolation
1990	Zenger, G.	sparse grids
2000	Stromberg, deVore	hyperbolic wavelets
2010	????	

Application areas include:

- quadrature
- interpolation
- data compression
- solution of PDEs
- integral equations
- eigenvalue problems

Basic principles of sparse grids

- 1-dim multilevel sequence of operators and spaces

$$P_l : V^{(1)} \to V_l \qquad V_l \qquad l \in \mathbb{N}$$

- Sequence of differences, telescopic approach

$$\Delta_l := (P_l - P_{l-1}) : V^{(1)} \to V_l \ominus V_{l-1} =: W_l$$

- d-dim. product construction $\mathbf{l} = (l_1, l_2, \ldots, l_d) \in \mathbb{N}^d$

$$\Delta_{\mathbf{l}} := \bigotimes_{j=1}^{d} \Delta_{l_j} = \bigotimes_{j=1}^{d} (P_{l_j} - P_{l_j - 1}) : V^{(d)} \to W_{\mathbf{l}} \qquad f_{\mathbf{l}} = \Delta_{\mathbf{l}}(f) \in W_{\mathbf{l}}$$

- Appropriate truncation of resulting multivariate expansion

$$\mathbb{N}^d \to \mathfrak{I} \subset \mathbb{N}^d$$

$$P = \sum_{\mathbf{l} \in N^d} \Delta_{\mathbf{l}} \;\to\; P_{\mathfrak{I}} = \sum_{\mathbf{l} \in \mathfrak{I}} \Delta_{\mathbf{l}}$$

Examples of multiscale expansions, 1d

- Integration: $P_l = Q_l : V^{(1)} \rightarrow V_l = \mathbb{R}$
  - Sequence of nested or non-nested point sets and weights, size: $n_l = l$ or $n_l = 2^l + 1$

    => various sparse grid quadrature rules
- Interpolation $P_l = I_l : V^{(1)} \rightarrow V_l$, approximation $P_l = A_l : V^{(1)} \rightarrow V_l$
  - Local piecewise polynomials, multiscale expansion: hierarchical basis, interpolets, wavelets, multilevel basis, size: $n_l = 2^l + 1$ $\qquad |W_l| = 2^{l-1}$

    => sparse grid finite element spaces
  - Global polynomials: Fourier series, Chebyshev, Legendre, Hermite, Bernoulli polynomials

    size $n_l = l$ or $n_l = 2^l + 1$ $\qquad |W_l| = 1$ or $|W_l| = 2^{l-1}$

    => total degree / hyperbolic cross approximation

Regular sparse grid approach

- Index sets

$$\mathfrak{I}_n^{\text{full}} = \left\{ \mathbf{l} \in \mathbb{N}^d : |\mathbf{l}|_\infty = \max_{j=1,\dots,d} l_j \le n \right\}$$

$$\mathfrak{I}_n^{\text{sparse}} = \left\{ \mathbf{l} \in \mathbb{N}^d : |\mathbf{l}|_1 = \sum_{j=1}^{d} l_j \le n + d - 1 \right\}$$

- The hierarchical representation is then

$$P_n^{\text{sparse}} = \sum_{|\mathbf{l}|_1 \le n+d-1} \Delta_{\mathbf{l}} \qquad P_n^{\text{sparse}}(f) = \sum_{|\mathbf{l}|_1 \le n+d-1} \Delta_{\mathbf{l}}(f)$$

- Other representations:
  - generating system
  - Lagrange system over SG points
  - semi-hierarchical
  - combination method

The combination technique

- A simple alternative representation is [G., Schneider, Zenger 91],

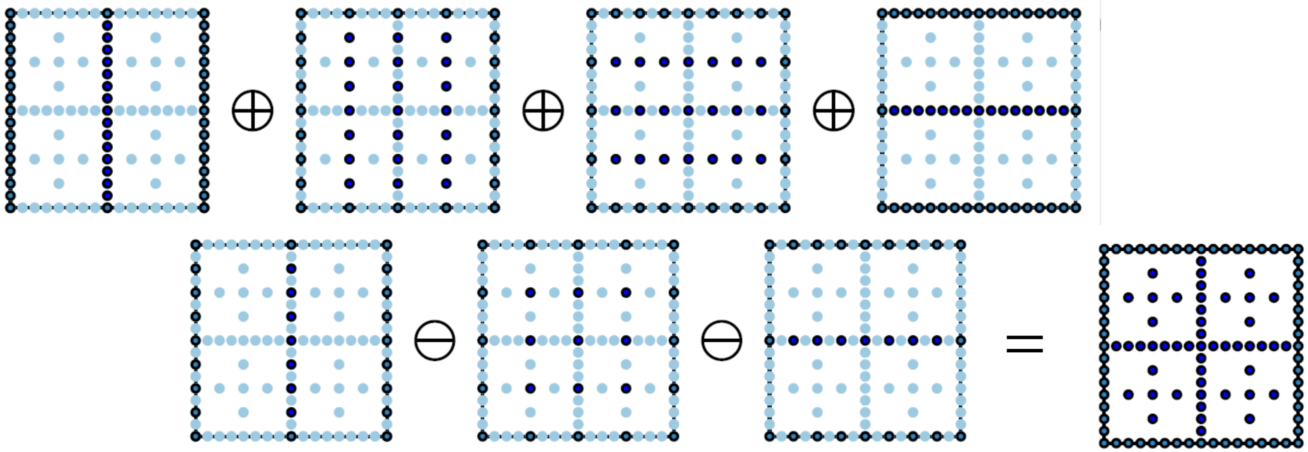
$$P_n^{\text{combi}} = \sum_{n \le |\mathbf{l}|_1 \le n+d-1} (-1)^{n+d-|\mathbf{l}|_1-1} \binom{d-1}{|\mathbf{l}|_1 - 1} P_{\mathbf{l}} \qquad P_{\mathbf{l}} := \bigotimes_{j=1}^{d} P_{l\,j}$$

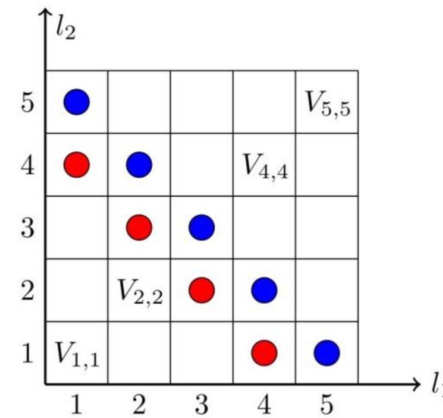
  – Involves just the (anisotropic) full grid discretizations $P_{\mathbf{l}}$ on different levels and linearly combines them

- 2D example

$$P_n^{\text{combi}} = \sum_{|\mathbf{l}|_1 = n+d-1} P_{\mathbf{l}} - \sum_{|\mathbf{l}|_1 = n+d-2} P_{\mathbf{l}}$$

$n = 4$





level indices, $n = 5$

The combination technique

- Redundant representation but allows the simple reuse of existing code
- Completely parallel computation of the subproblems $P_{\mathbf{l}}$
- Corresponds to a certain multivariate extrapolation method [Rüde 91]
- Necessary: Existence of a pointwise error expansion.
  - Euler-Maruyama of stochastic ODE: additive expansion (leading error term) of mean square error
- Multilevel-Monte Carlo is just 2-d combination method
  - Variance and bias for the two dimensions and a proper refinement rule which reflects the MC and the Euler-Maruyama rates [Gerstner12, Harbrecht,Peters,Siebenmorgen13]

A priori construction of sparse grids

- In general: Given
  - a class of functions and an error norm
  - an associated bound $b(\mathbf{l})$ for the benefit of $\Delta_{\mathbf{l}}$
  - a bound $c(\mathbf{l})$ for the cost of $\Delta_{\mathbf{l}}$
- We can a-priori derive a (quasi-) optimal sparse grid

  by solving a binary knapsack problem [Bungartz+G.03]

$$\max \sum_{\mathbf{l} \in N^d} \alpha_{\mathbf{l}} \cdot b(\mathbf{l}) \quad \text{such that} \quad \sum_{\mathbf{l} \in N^d} \alpha_{\mathbf{l}} \cdot c(\mathbf{l}) \leq C_{fix} \qquad \alpha_{\mathbf{l}} \in \{0,1\}$$

  and setting $\Im_C = \{ \mathbf{l} \in \mathbb{N}^d : \alpha_{\mathbf{l}} = 1 \}$

- Boils down to just sorting the quotients $b(\mathbf{l})/c(\mathbf{l})$ of the benefit versus cost according to its size and taking the largest indices into account

L^2-norm-based sparse grids in H^2_{mix}

- Representation $f(\mathbf{x}) = \sum_{\mathbf{l}} f_{\mathbf{l}}(\mathbf{x})$ $\quad f_{\mathbf{l}}(\mathbf{x}) \in W_{\mathbf{l}}$ $\quad \mathbf{x} = (x_1,..,x_d)$

  $\mathbf{l} = (l_1,..,l_d)$

- Cost per subspace $c(\mathbf{l}) = \dim(W_{\mathbf{l}}) = 2^{|\mathbf{l}-\mathbf{1}|_1}$
- Benefit for accuracy

$$\| f_{\mathbf{l}} \|_2 \le b(\mathbf{l}) = 3^{-d} \cdot 2^{-2|\mathbf{l}|_1} \cdot | f |_{\mathbf{2},mix} = O(2^{-2|\mathbf{l}|_1})$$

- Choice of best subspaces ? Knapsack problem !

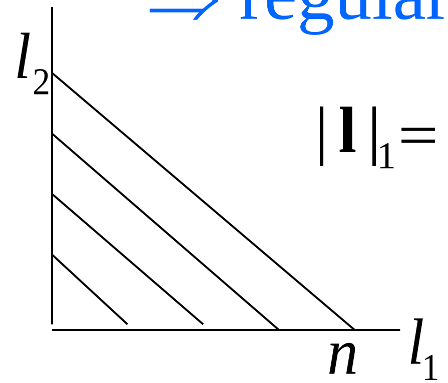
  => local benefit²/cost ratio

$$b^2(\mathbf{l}) / c(\mathbf{l}) \approx \frac{2^{-4\cdot|\mathbf{l}|_1}}{2^{|\mathbf{l}|_1}} = 2^{-5\cdot|\mathbf{l}|_1}$$

$$V_n^{(d,opt)} = \bigoplus_{|\mathbf{l}|_1 = n+d-1} W_{\mathbf{l}}$$

$\Rightarrow$ regular sparse grid space

$|\mathbf{l}|_1 = n + d - 1$ isoline



Anisotropic sparse grids

- Non-equal directions
  - Weighted Sobolev spaces [Sloan+Wozniakowski93]

$$H^{r}_{\gamma,mix}$$

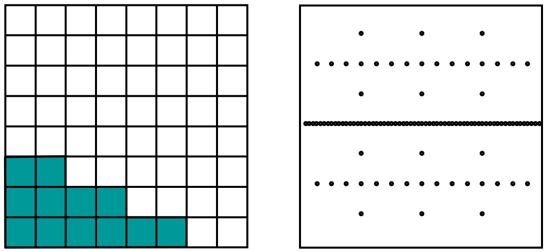
  - Anisotropic smoothness spaces [Gerstner+G. 98, G.+Zung15]

$$H^{s_1,s_2,\ldots,s_d}_{mix} = H^{s_1}(I_1) \otimes H^{s_2}(I_2) \otimes \cdots \otimes H^{s_d}(I_d)$$

  - Different dimensions for different directions [G.+Harbrecht 11]

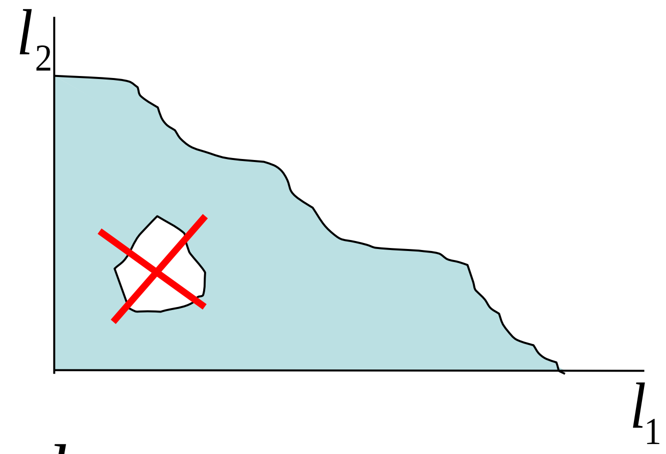
$$H^{s_1}(\Omega_1) \otimes H^{s_2}(\Omega_2) \otimes \ldots \otimes H^{s_d}(\Omega_d)$$

- Via knapsack problem:
  - A priori construction of optimal anisotropic sparse index sets
  - log-terms disappear



Generalized sparse grids

- General index sets $\mathfrak{I} \subset \mathbb{N}^d$



- Downward closed set, no holes

$$\mathbf{l} \in \mathfrak{I} \;\Rightarrow\; \mathbf{l} - e_j \in \mathfrak{I} \quad j = 1, \ldots, d$$

- Associated sparse grid operator $P_{\mathfrak{I}} = \sum_{\mathbf{l} \in \mathfrak{I}} \Delta_{\mathbf{l}}$

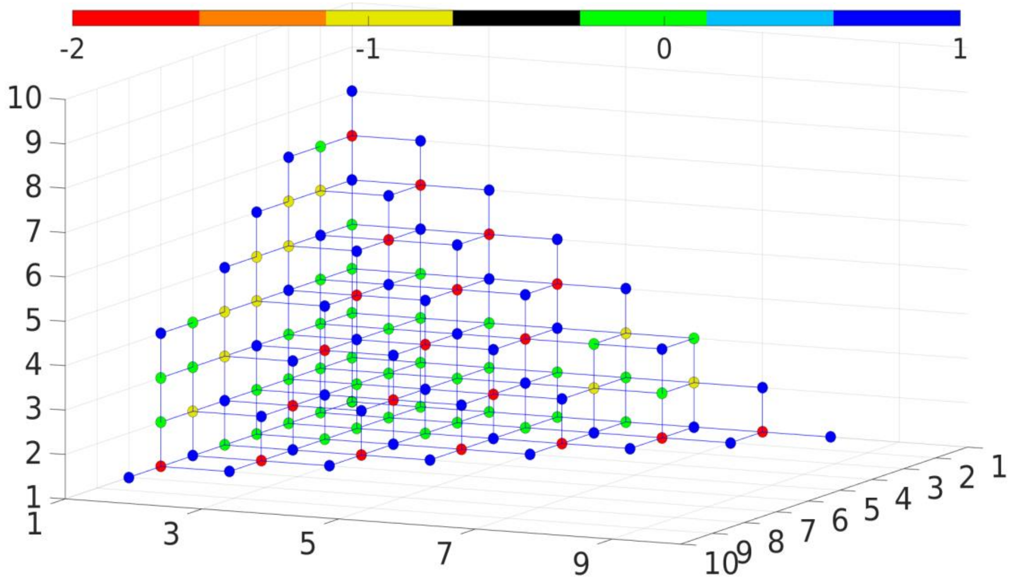
- Associated space and associated function

$$V_{\mathfrak{I}} = \bigoplus_{\mathbf{l} \in \mathfrak{I}} W_{\mathbf{l}} \qquad P_{\mathfrak{I}} f = \sum_{\mathbf{l} \in \mathfrak{I}} \Delta_{\mathbf{l}}(f) = \sum_{\mathbf{l} \in \mathfrak{I}} f_{\mathbf{l}}$$

The combination technique

- Can also be generalized to a given downward closed index set $\mathfrak{I}$

$$P_{\mathfrak{I}} = \sum_{\mathbf{l} \in \mathfrak{I}} c_{\mathbf{l}} P_{\mathbf{l}}$$



- Combination coefficient

$$c_{\mathbf{l}} = \sum_{\mathbf{z}=\mathbf{0}}^{\mathbf{1}} (-1)^{|\mathbf{z}|_1} \chi^{\mathfrak{I}}(\mathbf{l}+\mathbf{z})$$

with characteristic function $\chi^{\mathfrak{I}}$ on the index set $\mathfrak{I}$

- Again: just (anisotropic) full grid discretizations $P_{\mathbf{l}}$ on different levels get linearly combined
- Note: many coefficients on the lower levels are zero

Tensor product sparse grids

- Examples:
  - space $\times$ time, $d_1 = 3, d_2 = 1,$ parabolic problems
  - space $\times$ parameters $d_1 = 3, d_2 = 10 - 20$

    but smooth in parameter variables
  - space $\times$ stochastics $d_1 = 3, d_2 = \infty$

    but analytic in stochastic variables
- Main result: Curse of dimension only w.r.t. the larger dimension and/or the lower smoothness

  [G.+Harbrecht11], [G.+Zung15]
- Time, parametrization and stochastic coordinates disappear in the overall complexity rate

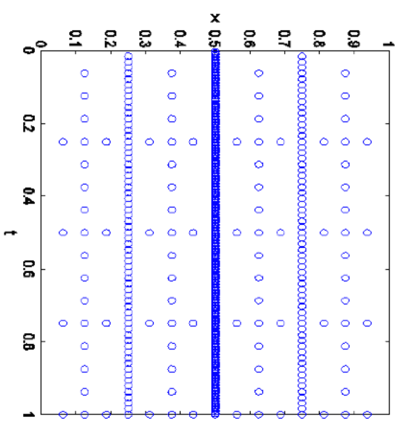
=> just space discretization matters

Sparse space-time grids

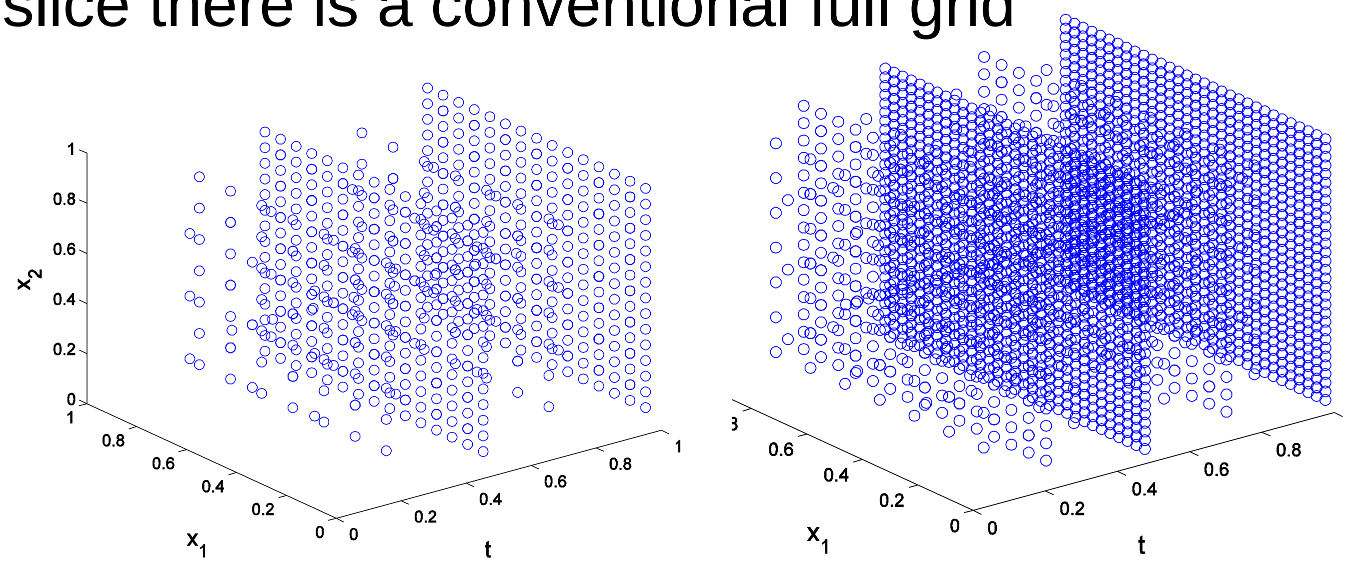
- Approximation error and necessary regularity [G.+Oeltz07]

$$\inf_{u_n \in V_n^0} \left\| u - u_n \right\|_{H^1(\Omega) \otimes L^2(0,T)} \leq c\; 2^{-n} \left\| u \right\|_{H^2(\Omega) \otimes H^2((0,T))}$$

  - Classical regularity theory (Ladyzenskaja, Wloka) $u \in H^2(\Omega) \otimes H^2((0,T))$
  - Sparse space-time grids posses same approximation rate as full space-time grids but just cost complexity of space problem
  - In each time slice there is a conventional full grid



space dimension 1, space-time sparse grid, Euler case



space dimension 2, space-time sparse grid, Cranck-Nicolson case, n=4,5:

Stochastic and parametric PDEs

- Solutions of stochastic/parametric PDEs

$$\partial_t u(t,\mathbf{x},\mathbf{y}) - \nabla \cdot A(\mathbf{x},\mathbf{y}) \nabla f(t,\mathbf{x},\mathbf{y}) = r(t,\mathbf{x},\mathbf{y})$$

  live on product $(t,\mathbf{x},\mathbf{y}) \in T \times \mathbf{X} \times \mathbf{Y}$
  - of temporal domain T
  - of spatial domain $\mathbf{X}$ with $d_1 = 1,2,3$
  - and stochastic/parametric domain $\mathbf{Y}$ with d_2 large or even infinity.
- Often: Very high smoothness in $\mathbf{y}$-part
  - Here: especially weighted analyticity for the different coordinates, decay in covariance [Cohen,Devore,Schwab10,11]
  - Then, even infinite-dimensional $\mathbf{Y}$ become treatable
- Sparse grid not only within stochastics but also between spatial, temporal and stochastic domain

Sparse grids and analytic functions

- Analytic regularity in polydisc with radii $\mathbf{r} := (r_1, \ldots r_d)$
- Sequence of smoothness indices $\mathbf{a} = (a_1, \ldots a_d) = \log(\mathbf{r})$
- With global polynomials: $|\Delta_{\mathbf{k}}(f)| \leq c \cdot e^{-(a_1 k_1 + \ldots + a_d k_d)}$
- Accuracy with respect to the involved #dof M

  [Beck,Nobile,Tamellini,Tempone12,14], [Tran,Webster,Zhang15], [G.+Oettershagen15]

$$gm(\mathbf{a}) = \left(\prod_{j=1}^{d} a_j \right)^{1/d} \qquad \kappa(d) = (d\,!)^{1/d} > d/e \qquad O(e^{-gm(\mathbf{a})\kappa(d)M^{1/d}} M^{(d-1)/d})$$

- For the infinite-dimensional case:
  - Logarithmic growth => algebraic rate

    [Todor,Schwab07], [Cohen,Devore,Schwab10,11]

$$\beta > 1 \qquad \sum_{j=1}^{\infty} \frac{1}{e^{a_j/\beta} - 1} < \infty \qquad O(M^{-(\beta-1)})$$

Stechkin´s Lemma

  - Linear growth => subexponential rate [G.+Oettershagen15], [Tran,Webster,Zhang15]

$$\alpha > 0 \qquad a_j \geq \alpha \cdot j \qquad O(M^{-\frac{3}{8}\alpha \cdot \sqrt{\log(M)}} M^{1+\frac{\alpha}{4}} \log(M)^{-1/2})$$

Stechkin´s Lemma can not show this rate but gives only an algebraic bound

Dimension-adapted sparse grids

- So far: function class known,
  - a-priori choice of best subspaces by optimization
  - size of benefit/cost ratio indicated if subspace is relevant

    => sparse grid patterns for $\mathfrak{I}$
- Now: for given single function f
  - adaptively build up a set $\mathfrak{I}$ of active indices
  - benefit $b(\mathbf{l}) := \| \Delta_{\mathbf{l}}(f) \|^2$, i.e. local error-indicator of f
  - cost $c(\mathbf{l}) = |W_{\mathbf{l}}|$ for subspace $W_{\mathbf{l}}$,
  - benefit/cost indicator $\varepsilon(\mathbf{l}) := b(\mathbf{l}) / c(\mathbf{l})$
  - refinement strategy to build new index set,
  - global stopping criterion => sparse grid pattern $\mathfrak{I}$
- Directions $T \times \mathbf{X} \times \mathbf{Y}$ with product of different smoothness

The adaptive combination algorithm

Result: Solution $\boldsymbol{u}^c$ with error $< \mathrm{TOL}$.

$\boldsymbol{l} := (1, \ldots, 1)$;

$A := \{\boldsymbol{l}\}$; /* active index set */

$O := \emptyset$; /* old index set */

$\epsilon_{\boldsymbol{l}}$; /* local benefit/cost indicator */

E ; /* global error indicator */

while $E > TOL$ **do**

  select $\boldsymbol{l} \in A$ with largest $\epsilon_{\boldsymbol{l}}$;

  $O = O \cup \{\boldsymbol{l}\}, \quad A = A \backslash \{\boldsymbol{l}\}$;

  for $t \leftarrow 1$ **to** d **do**

    $\boldsymbol{j} = \boldsymbol{l} + \boldsymbol{e}_t$;

    if $\boldsymbol{j} - \boldsymbol{e}_k \in O \quad \forall\, k = 1, \ldots, d$ **then**

      $A = A \cup \{\boldsymbol{j}\}$;

      Solve problem with level-parameters $\boldsymbol{j}$;

      Compute local benefit/cost indicator for $\boldsymbol{j}$;

    end

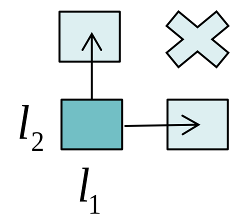
  end

  Compute new global error indicator E;

end

Compute $\boldsymbol{u}^c$ on index set $\mathcal{I} = O \cup A$;

refinement rule



downward closedness

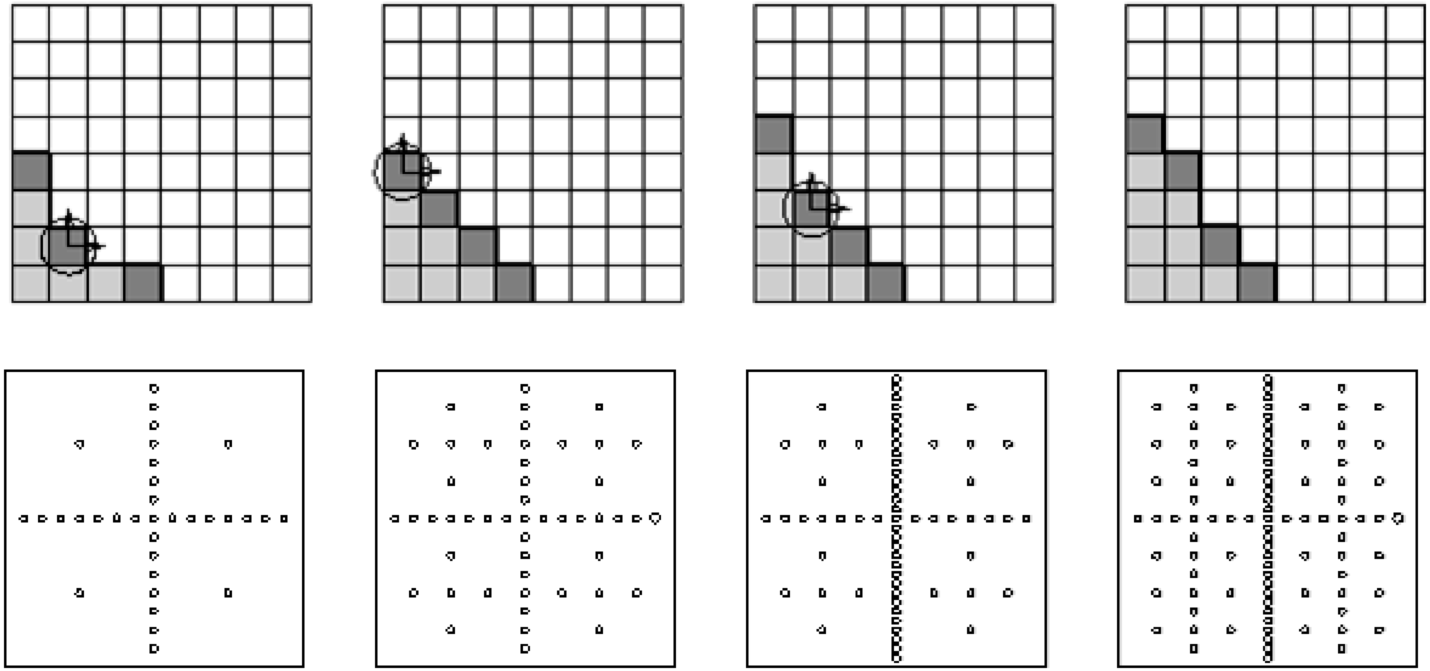
simple extension to dimension-adaptive version exists => UQ14

Example

- Evolution of the algorithm:

index sets:

corresponding grids:



- As any adaptive heuristics: may terminate too early
- If mixed regularity not present, refinement to the usual full grid

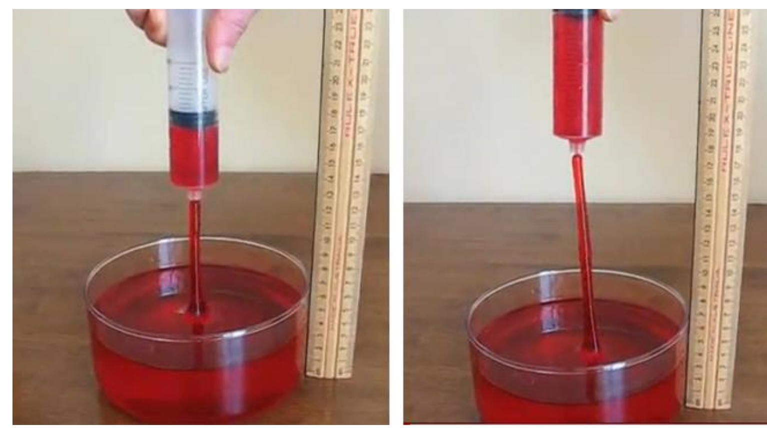
Application: Non-Newtonian fluids

- Classical Newtonian fluids: Obey Newton´s law of viscosity, stress tensor is proportional to load/force

- But various complex fluids show strange behavior which is not correctly described

Barus effect

Weissenberg effect



tubeless siphon effect

Application: Non-Newtonian fluids

- Non-Newtonian fluids contain microstructures which are the reason for their unusual properties
  - Examples: paint, toothpaste, shampoo, blood, oils
- Polymeric fluids are a subset of non-Newtonian fluids
  - Long-chained molecules in a Newtonian solvent
  - Viscoelasticity due to interaction of elastic molecules and drag forces in basic flow
- A macroscopic model like the Navier Stokes equations + macrosopic extensions is no longer sufficient
- Needs to be augmented by model on the micro scale

=> Two scale modelling

Mathematical modelling

- The conservation equations for polymeric fluids are the same as for the Newtonian case, but the presence of polymer molecules contributes a polymeric extra-stress tensor $\boldsymbol{\tau}_p$ and an additional polymeric viscosity η_p such that the viscosity ratio $\beta < 1$
- The Navier-Stokes equations are now

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = \frac{1}{\mathrm{Re}} \beta \Delta \mathbf{u} - \nabla p + \frac{1}{\mathrm{Re}} \nabla \cdot \boldsymbol{\tau}_p$$

conservation of momentum

$$\nabla \cdot \mathbf{u} = 0$$

+ b.c., with Reynolds number Re

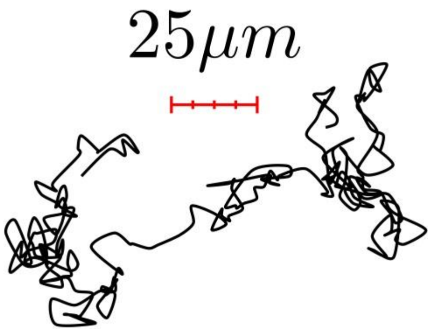
and viscosity ratio $\beta = \dfrac{\eta_s}{\eta_s + \eta_p}$

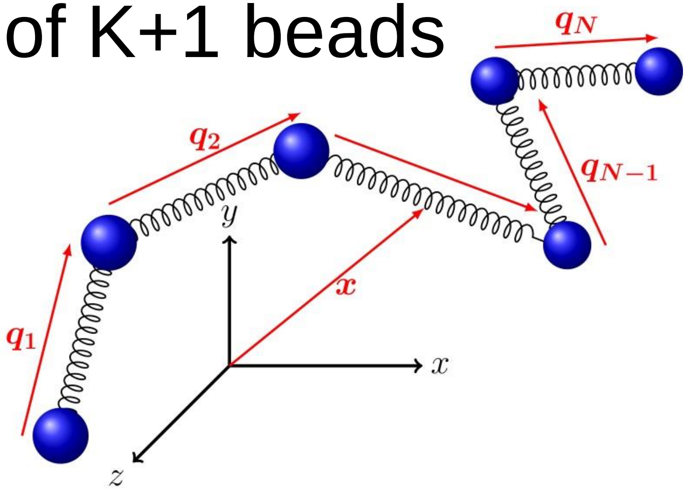
η_s solvent viscosity

η_p polymeric viscosity

Microscopic modelling

- On the microsocopic scale, a polymer chain is modelled by a spring chain of K+1 beads





- Position $\mathbf{x}$ in physical space/flow domain $\Omega \subset \mathbb{R}^3$
- Orientations $\mathbf{q}_1, \ldots, \mathbf{q}_K$ in configuration space $\Gamma \subset \mathbb{R}^{3K}$
- Probability to find chains at time t with position in .$[\mathbf{x}, \mathbf{x} + d\mathbf{x}]$ and orientations in $[\mathbf{q}_1, \mathbf{q}_1 + d\mathbf{q}_1] \ldots [\mathbf{q}_K, \mathbf{q}_K + d\mathbf{q}_K]$

$$\psi : \Omega \times \Gamma \times [0, T] \rightarrow \mathbb{R}^+, \; (\mathbf{x}, \mathbf{q}_1, \ldots, \mathbf{q}_K, t) \rightarrow \psi\,(\mathbf{x}, \mathbf{q}_1, \ldots, \mathbf{q}_K, t)$$

Fokker-Planck equation

- The function ψ is a pdf, i.e. $\psi \geq 0, \; \int_\Gamma \psi = 1$
- The application of Newton´s 2nd law to the forces acting on chain leads to the Fokker-Planck equation

$$\frac{\partial \psi}{\partial t} + \nabla_{\mathbf{x}} \cdot (\mathbf{u}\psi) + \sum_{i=1}^{K} \nabla_{\mathbf{q}_i} \cdot \left((\nabla_{\mathbf{x}} \mathbf{u})^T \mathbf{q}_i \, \psi - \frac{1}{4De} \sum_{j=1}^{K} A_{ij} \mathbf{F}(\mathbf{q}_j) \, \psi \right) = \frac{1}{4De} \sum_{i=1}^{K} \sum_{j=1}^{K} A_{ij} \nabla_{\mathbf{q}_i} \cdot \nabla_{\mathbf{q}_j} \psi$$

Deborah number → (De)

with Rouse matrix $A = [-1 \; 2 \, -1]_K$

- Describes evolution of ψ under chain´s spring forces

$$\mathbf{F}(\mathbf{q}_1), .., \mathbf{F}(\mathbf{q}_K)$$

- Various models for spring force: Hooke: $\mathbf{F}(\mathbf{q}) = \mathbf{q}$

FENE: $\mathbf{F}(\mathbf{q}) = \dfrac{\mathbf{q}}{1 - \|\mathbf{q}\|^2 / b}, \quad \| q \|^2 \leq b,$ FENE-P: $\mathbf{F}(\mathbf{q}) = \dfrac{\mathbf{q}}{1 - \langle \mathbf{q}^2 \rangle / b}, \quad \langle \mathbf{q}^2 \rangle \leq b$

Coupling to the macro scale

- ψ represents polymeric configurations of micro-system
- Expectation in configuration space

$$\langle \cdot \rangle = \int_{\Gamma} \cdot \ \psi \ d\ \mathbf{q}_1 ... d\ \mathbf{q}_K$$

- Coupling of internal configurations of micro system to macroscopic stress tensor via Kramer´s expression

$$\boldsymbol{\tau}_p = C \sum_{i=1}^{K} \left(\langle \mathbf{q}_i \otimes \mathbf{F}(\mathbf{q}_i) \rangle - \mathbf{Id} \right)$$

Constant C depends on model, Deborah number, viscosity ratio

- Issues with the Fokker-Planck equation
  - becomes more singular for higher values of De [Suli, Knezevic08]

    => extremely fine numerical resolution needed [Lozinski, Owen 03]
  - $3+3K=3(K+1)$ -dimensional + time-dependent => curse of dim.

Stochastic microscopic modelling

- There is a formal equivalence between the Fokker-Planck equation and stochastic partial differential eq.

$$d\vec{\mathbf{Q}}(\mathbf{x},t) = \Big(-(\mathbf{u}\cdot\nabla)\vec{\mathbf{Q}}(\mathbf{x},t) + (\nabla\mathbf{u})\cdot\vec{\mathbf{Q}}(\mathbf{x},t) -$$

Deborah number →

$$- \frac{1}{4De} \; A\,\mathbf{F}(\vec{Q}(\mathbf{x},t)) \Big) dt + \sqrt{\frac{1}{2De}}\, d\vec{\mathbf{U}}(t)$$

  - Describes evolution of K random fields $\vec{\mathbf{Q}} = (\mathbf{Q}_1,\ldots,\mathbf{Q}_K)^T$ that represent the configuration vector $\vec{\mathbf{q}} = (\mathbf{q}_1,\ldots,\mathbf{q}_K)^T$
  - Brownian forces on the beads are modelled by the 3-dim. Wiener processes $\mathbf{W}_i(t),\; i = 1,\ldots,K+1$
  - The vector $\vec{\mathrm{U}}(t)$ consists of the component-wise differences

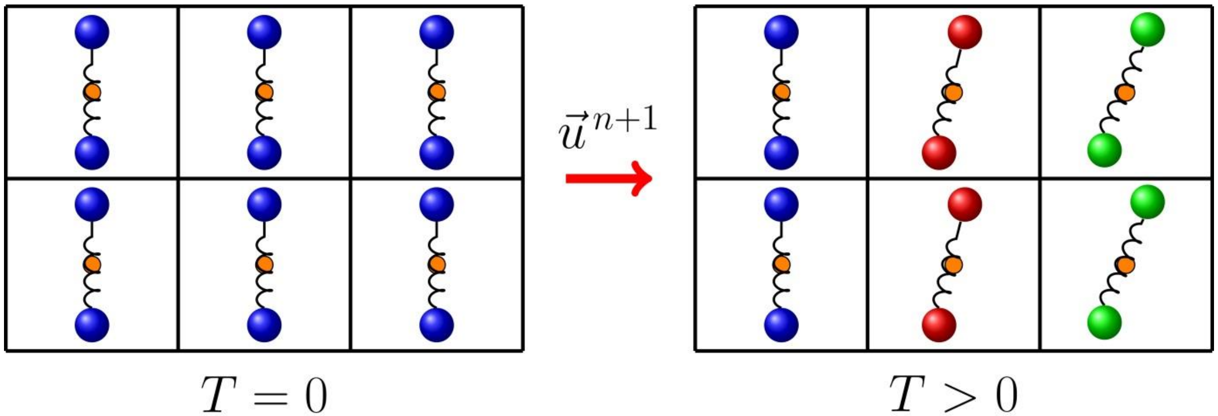
$$(\vec{\mathrm{U}}(t))_i = \mathbf{W}_{i+1}(t) - \mathbf{W}_i(t), \;\; i = 1,\ldots,K$$

Stochastic microscopic simulation

- Brownian configuration fields (BCF) [Hulsen97]

  Random field $\vec{\mathbf{Q}}(\mathbf{x},t)$ for configuration

- Discretization of x-space: the M_G grid cells make from the parabolic SPDE a system of SODEs (MoL)
- Discretization of SODE-system: Put M_B configuration fields in each of the M_G space grid cells and evolve their configuration discretely over time, i.e. all $M_G \cdot M_B$ configuration fields have fixed spatial positions (Eulerian view).



Stochastic microscopic simulation

- In each grid cell $k = 1, \ldots, M_G$ with center $\mathbf{x}_k$ we solve/ integrate the stochastic DE for a number M_B of stochastic realizations $\vec{\mathbf{Q}}^{(j)}(\mathbf{x}_k, t),\ \ j = 1, \ldots, M_B$
- They are distributed according to the known equilibrium density ψ for $t = 0$
- But we do not know ψ for $t > 0$. Thus, we approximate the first moments $\langle \mathbf{Q}_i(\mathbf{x}_k, t) \otimes \mathbf{F}(\mathbf{Q}_i(\mathbf{x}_k, t)) \rangle$ in Kramer´s relation as

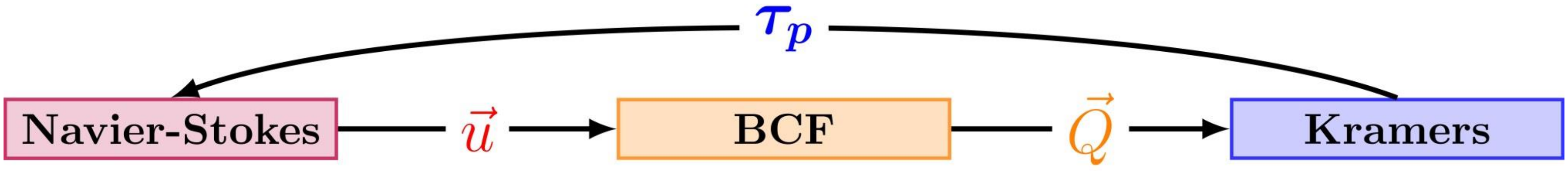
$$\boldsymbol{\tau}_p(\mathbf{x}_k, t) = C \sum_{i=1}^{K} \left(\langle \mathbf{Q}_i(\mathbf{x}_k, t) \otimes \mathbf{F}(\mathbf{Q}_i(\mathbf{x}_k, t)) \rangle - \mathbf{Id} \right)$$

$$\approx C \sum_{i=1}^{K} \left(\frac{1}{M_B} \sum_{j=1}^{M_B} \mathbf{Q}_i^{(j)}(\mathbf{x}_k, t) \otimes \mathbf{F}(\mathbf{Q}_i^{(j)}(\mathbf{x}_k, t)) - \mathbf{Id} \right)$$

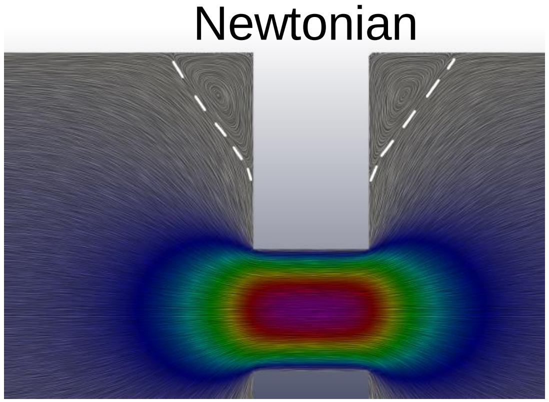
i.e. we replace the integral by Monte Carlo quadrature

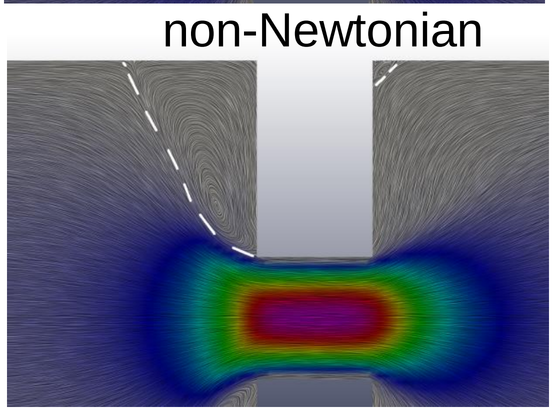
Numerics

- Navier Stokes equations:
  - Uniform grid cells, staggered grid, cell centers p, $\boldsymbol{\tau}_p$, cell faces $\mathbf{u}$
  - WENO for convective terms, 2nd order scheme for other terms
  - Euler or Crank-Nicolson in time, CFL-condition
  - Chorin-like projection method
- Microscale stochastic equations:
  - M_B stochastic samples for each grid cell => $M_G \cdot M_B$ samples
  - QUICK for convective terms
  - Explicit Euler-Maruyama, semi-implicit Euler for FENE
  - Same time step size as for NS equations
  - Variance reduction scheme with equilibrium control variates



Issues





- Code works as expected
- But: Huge memory requirements and huge computing times due to large number M_B of realizations in each cell
- Example for 3D multi-scale problem
  - Flow domain Ω with
    - M_G = 100x100x100 grid cells
    - M_B = 10.000 stochastic realizations in each grid cell
  - Total memory requirements:
    - 8 MB for the pressure field p
    - 24 MB for the velocity field $\mathbf{u}$
    - 48 MB for the six independent components of $\boldsymbol{\tau}_p$
    - 75 GB*N for all the $M_G \cdot M_B$ stochastic variables
  - Some months of computing time

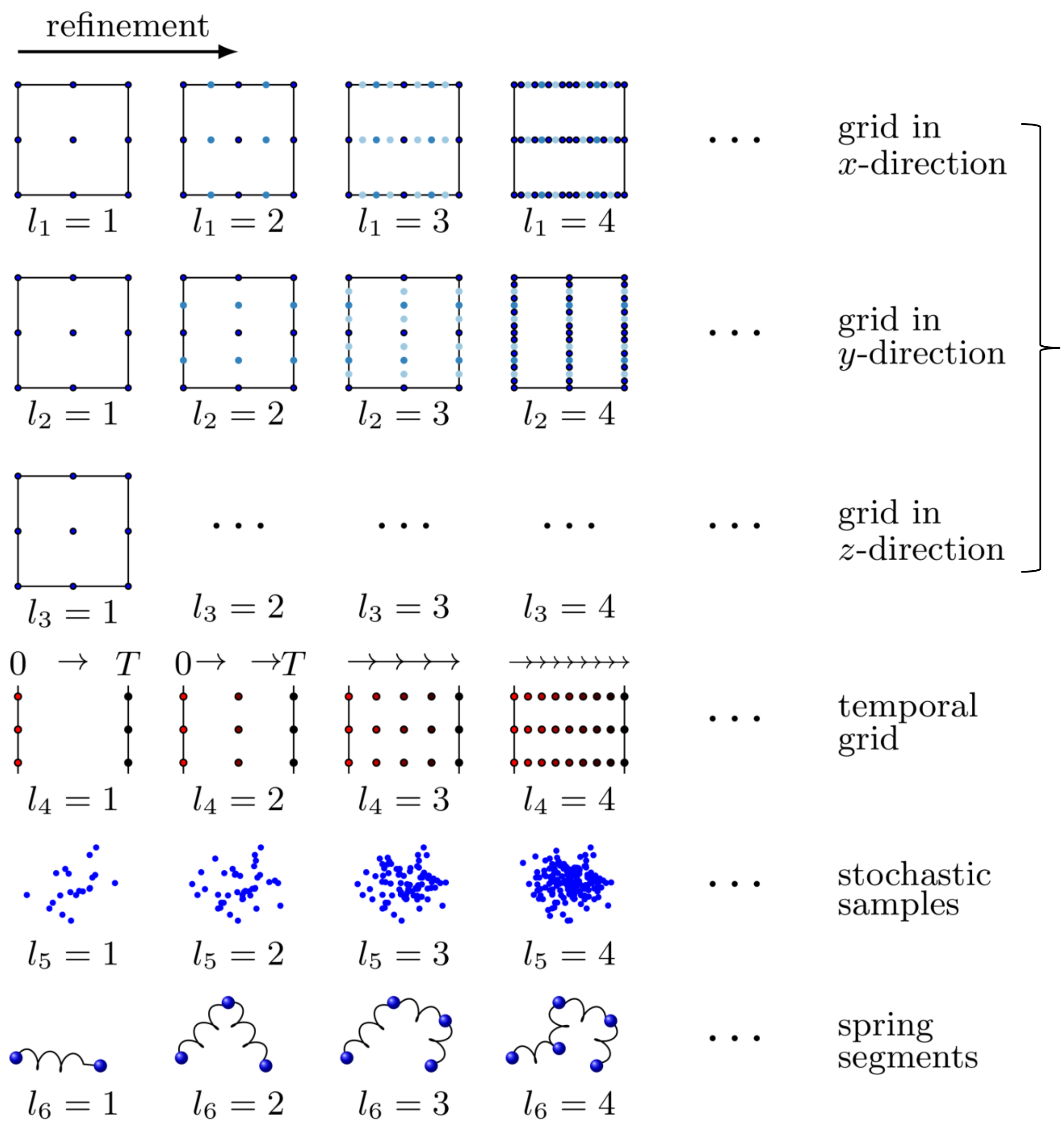
Sparse grid approach

- Consider our multiscale flow problem in more detail.
- We have the problem parameters:

  mesh width, time step size, stochastic realizations, springs
- How can we improve on computational complexity ?
  - Instead of MC use QMC
  - Multilevel-MC, MLQMC for stochastic ODEs (time + stoch.)

    This is just a certain 2d combination technique/
    sparse grid approach [Gerstner 12] [Harbrecht,Peters,Siebenmorgen13]
  - Combination technique in all 3 discretization parameters
    i.e. for space x time x stochastics,
    and for model parameter K, i.e. …. x number of springs
  - If the optimal combination formula is not a priori known:
    run the (dimension)-adaptive algorithm

Coordinates for the combination method



we use only an isotropic grid in our NS solver

Indicators for the combination method

- Approximation of the vector $\mathbf{u}$ and the tensor $\boldsymbol{\tau}_p$
- Compute benefits $b(\mathbf{l})$ and costs $c(\mathbf{l})$ componentwise
- One index set for all components

- Weighted and scaled benefit/cost indicator

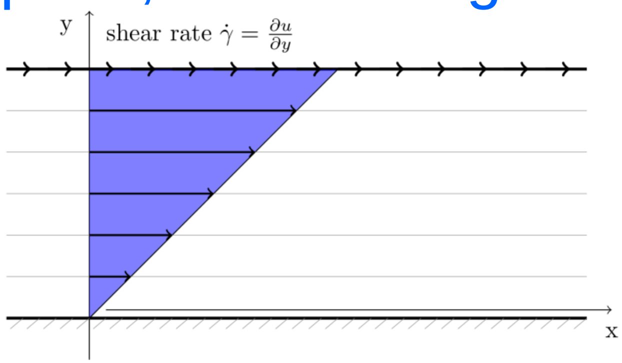
$$\varepsilon(\mathbf{l}) = \max\left\{ \omega \cdot \frac{\| b(\mathbf{l})(\mathbf{u}) \|_{2,2}}{c(\mathbf{l})(\mathbf{u}) \cdot \| b(1)(\mathbf{u}) \|_{2,2}}, (1-\omega) \cdot \frac{\| b(\mathbf{l})(\tau_p) \|_{F,2}}{c(\mathbf{l}) \cdot \| b(1)(\tau_p) \|_{F,2}} \right\}$$

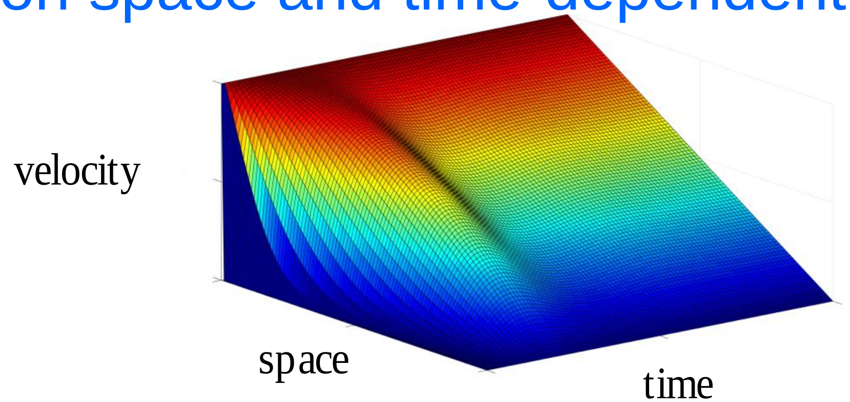
Scaling with initial level $b(1)$ not necessary if $\omega = 0$ or $\omega = 1$

Example 1: Couette flow

- Non-Newtonian fluid in a 2D channel.
  - Fluid is at rest at initial time t = 0, $De = 0.5$
  - Shearing of fluid over time with rate $\dot{\gamma} = du / dy$
  - Linear spring force model (dumbbell, K=1)
  - Probability density function $\psi : (x, \mathbf{q}, t) \in \mathbb{R}^4 \rightarrow \psi\ (x, \mathbf{q}, t) \in \mathbb{R}$

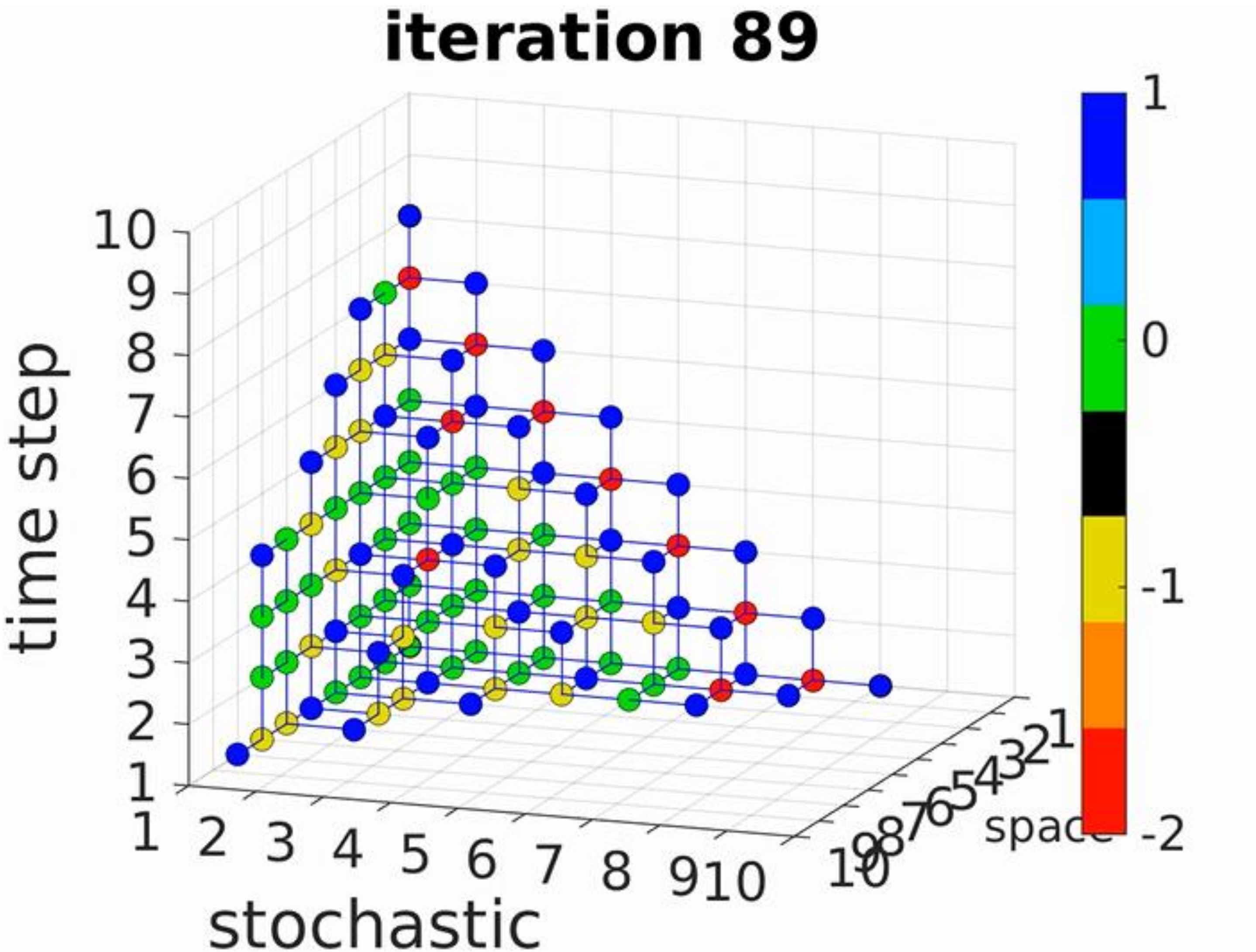
1d in space, 2d in configuration space and time-dependent





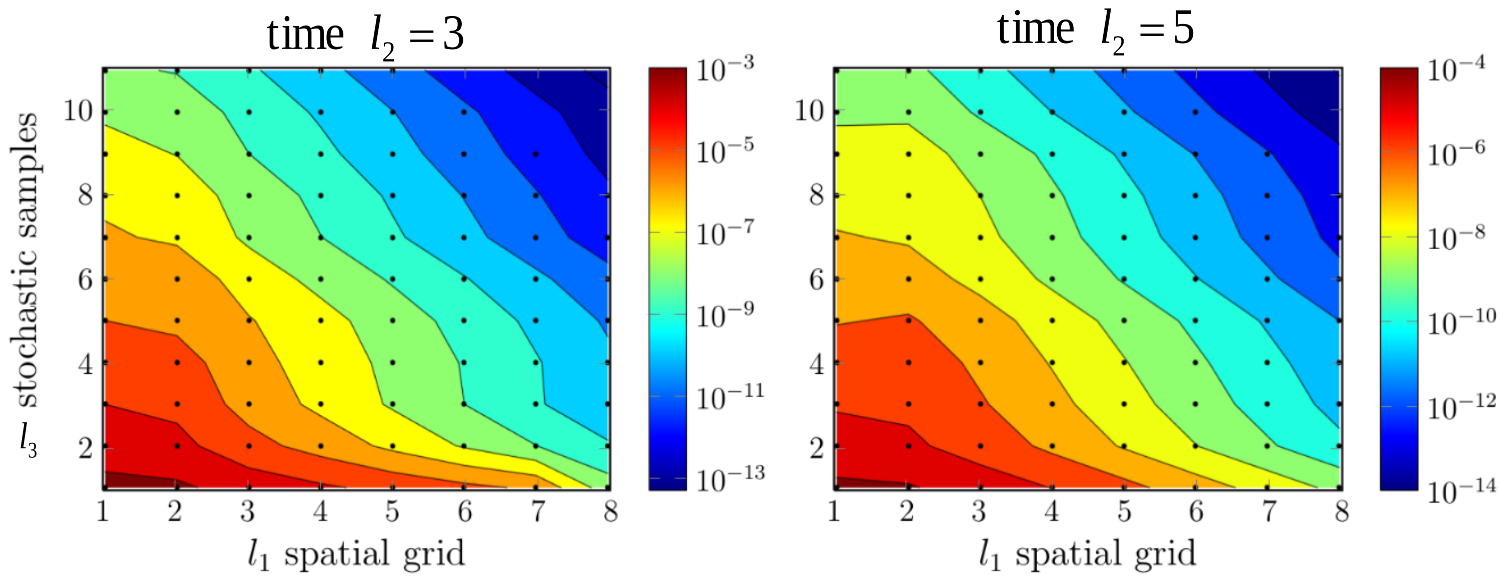
- Discretization:
  - Initial level $(1/\Delta x,\ 1/\Delta t,\ \text{samples}) = (4,\ 16,\ 256)$
  - Refinement from level to level by factor *2
  - Error indicator $\omega = 1$, we are after error in $\mathbf{u}$

Example 1 Couette flow

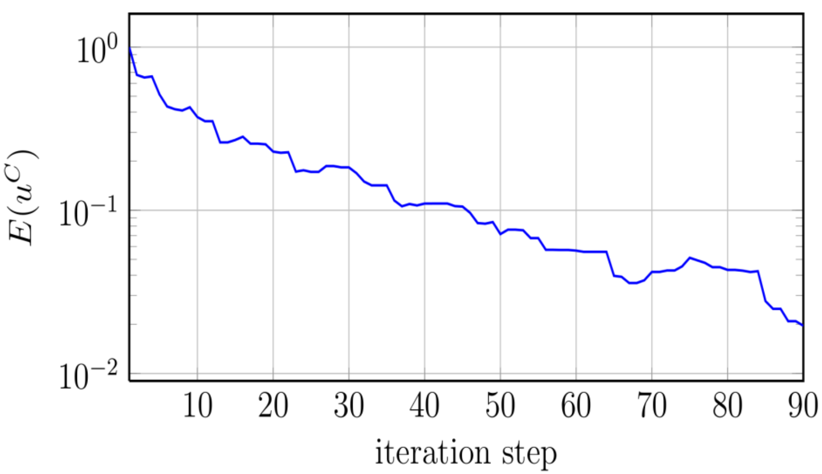


- Behaviour of adaptive combination technique

Example 1 Couette flow



- We asymptotically observe an anisotropic sparse grid structure

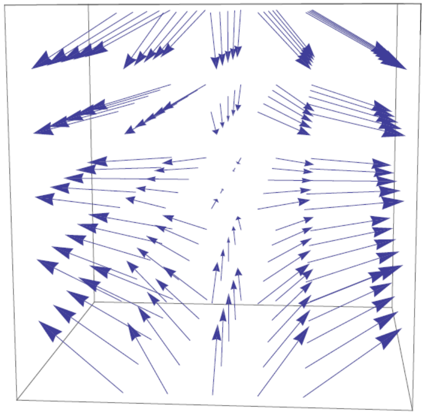


- Relative L_2 error of $\mathbf{u}_1$

- Comparison:
  - Full grid error $E(u_{6,6,6}) \approx 0.04$
    $E(u_{7,7,7}) \approx 0.01$
  - Cost (dof)
    full grid $C(u_{6,6,6}) \approx 5.4 \times 10^8$
    $C(u_{7,7,7}) \approx 4.3 \times 10^9$
    sparse grid $C(u^C) \approx 4.6 \times 10^7$

Example 2: Steady extensional flow

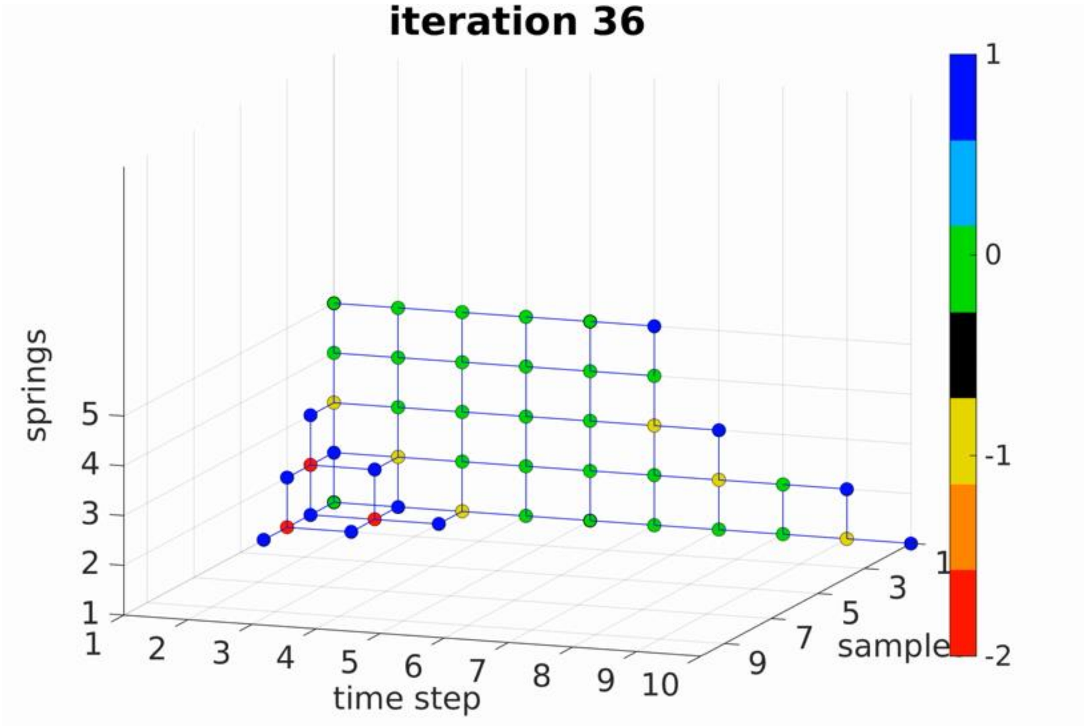
$\mathbf{u} = (\dot{\varepsilon}\, x, -\frac{\dot{\varepsilon}}{2} y, -\frac{\dot{\varepsilon}}{2} z)$

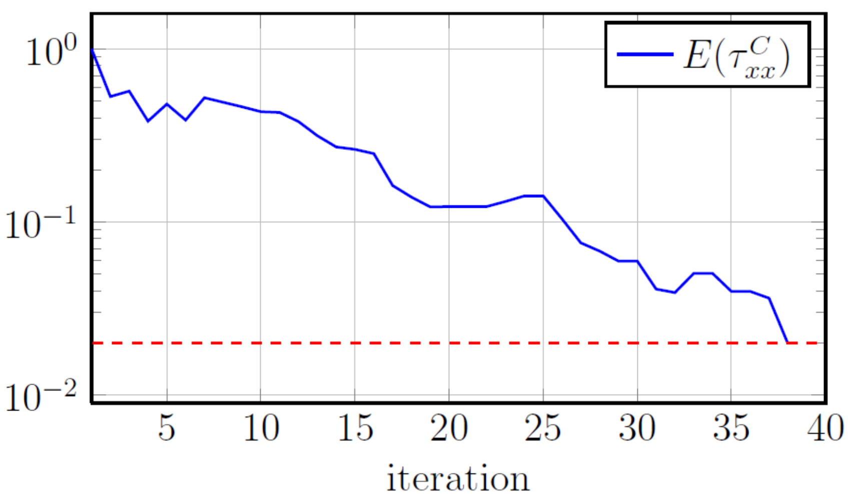


- Non-Newtonian fluid in a 3D domain.
  - Steady uniaxial extensional flow, $De = 1.0$
  - Stress tensor $\boldsymbol{\tau}_p$ is aimed for
  - FENE force model, K-spring chain
  - We vary the number K of springs up to 5
  - Probability density function $\psi : (\mathbf{q}, t) \in \mathbb{R}^{3K} \times \mathbb{R} \to \psi\,(\mathbf{q}, t) \in \mathbb{R}$

  3N-dimensional in configuration, time-dependent, number of springs, no space
- Discretization
  - Initial level $(\text{samples},\ 1/\Delta t,\ \text{springs}) = (1024,\ 2,\ 1)$
  - Refinement for time and samples from level to level by factor *2, refinement for springs by +1
  - Error indicator $\omega = 0$, we are after error in $\boldsymbol{\tau}_p$

Example 2: Steady extensional flow





Behaviour of adaptive combination technique

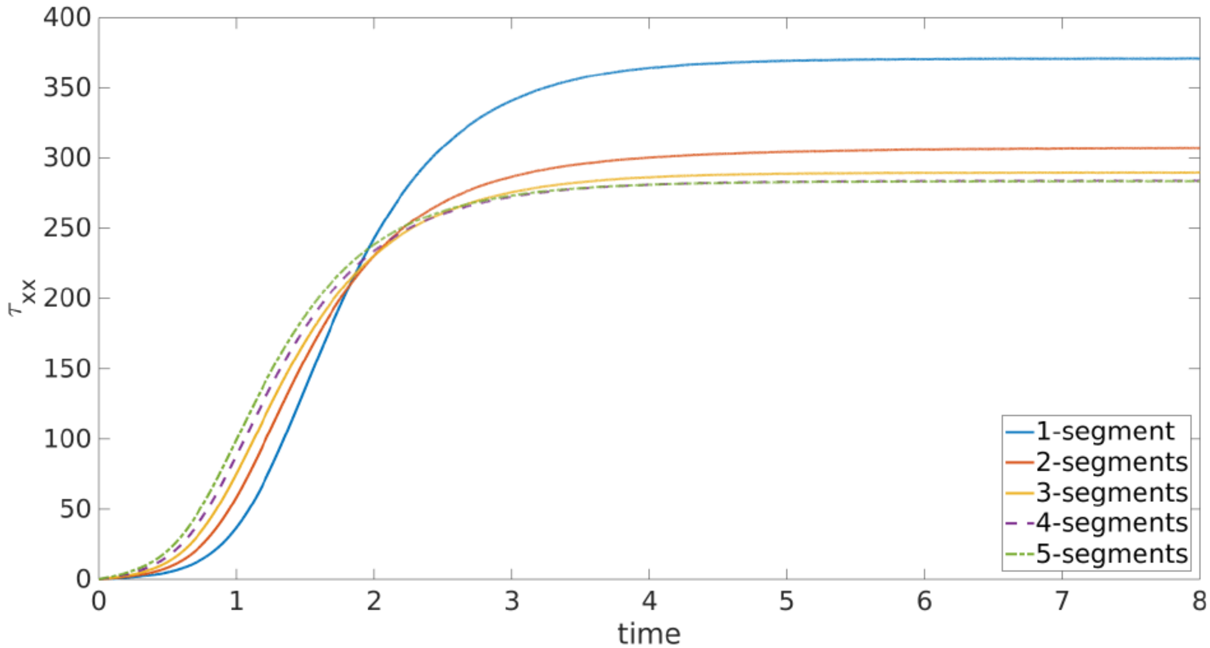
We observe:

- a sparse grid structure for all indices
- plus a nearly full grid between time and springs for the smallest sample size
- Different refinement: *2 versus +1

- Relative L_2 error for τ_{xx} of adaptive combination technique

Example 2: Steady extensional flow

- Convergence of model for rising number K of springs



- All results are computed on fine level with 2 million samples.
- Fixed stochastic time step width $\Delta t = 1/2048$

Concluding remarks

- Basic principles of sparse grids
- Optimization by knapsack problem
- Dimension-adaptive combination method
  - Solution of subproblems $P_{\mathbf{l}}$ on levels $\mathbf{l}$
  - Sparse grid approximation by linear combination
  - Refinement with hierarchical contributions $\Delta_{\mathbf{l}}$ and local cost
- Application to non-Newtonian flow
  - Two-scale problem, stochastic microscale
- Adaptive combination method works on discretization directions (space x time x samples) and also for model parameters (… x springs)

=> Allows to couple discretization and modelling errors

The C library HCFFT G.+Hamaekers

- Hierarchical sparse grid interpolation based on:
  - Fast Fourier transform (FFT), fast Sine and Cosine transform
  - Fast Chebyshev transform, Fast Legendre transform
  - Various other polynomial transforms
- Different hierarchical bases for different dimensions
- Dyadic and arbitrary, non-dyadic refined grids
- Several types of general sparse grids
- Dimension-adaptive sparse grids
- For high precision: possible use of long double
- Freely available at

www.hcfft.org

The flow solver

- Code NAST3DGPF which is freely available at

  http://www.nast3dgpf.de/