150 Jahre culture of excellence

Challenges for numerical analysis in large-scale simulations

S. Bauer*, D. Drzisga*, T. Horger*, M. Huber*, M. Mohr*,
U. Rüde*, L. Wunderlich*, B. Wohlmuth*
(LMU, FAU, TUM)

supported by the Gauss Centre for Supercomputing (GCS) and the DFG priority programme:
Software for Exascale Computing (1648)

Technical University of Munich (TUM)

July 11, Portland

2018 SIAM Annual Meeting

Overview

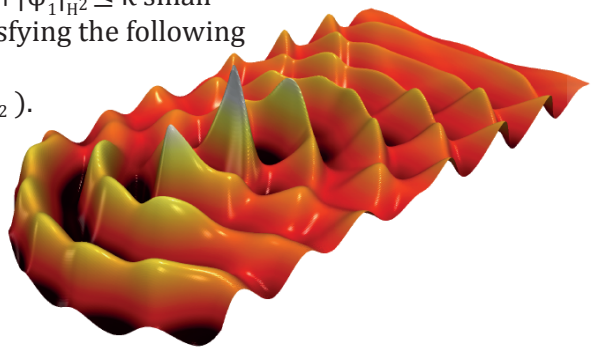
- **Aspects of simulation technologies for PDEs**
- **Hybrid assembling based on domain partitioning**
  - Surrogate polynomials for large scale FE
  - Local static condensation for patch-wise IGA
- **Large FE scale simulation**
  - All-at-once multigrid solver
  - Agglomeration for the coarse solver
- **Error estimation and control**
  - Adaptive error control for resilience
  - Adaptivity in sampling and surrogates

Aspects of modern simulation technologies for PDEs

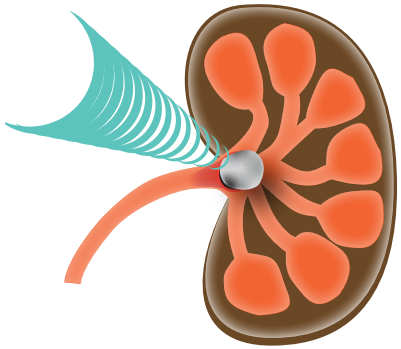
Mathematical modelling
The analysis challenge

Local well-posedness of Blackstock wave equation: Let Ω be an $C^{2,1}$ regular domain and c^2, $b > 0$, $k \in \mathbb{R}$. Assume that $|\psi_0|^2_{H^3} + |\psi_1|^2_{H^2} \leq \kappa$ small enough. Then there exists a unique solution $\psi \in X$, satisfying the following energy estimate

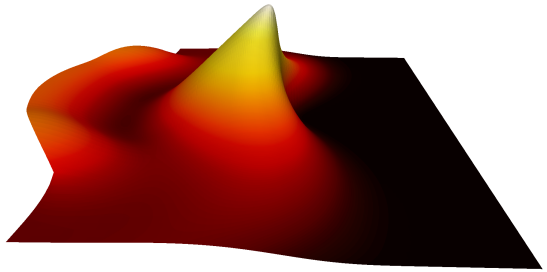
$$\|\psi\|^2_{L_\infty H^3} + \|\psi_t\|^2_{L_\infty H^2} \leq C(|\psi_0|^2_{H^3} + |\psi_1|^2_{H^2}).$$



Numerical Analysis
The algorithmic challenge



$$dJ(\Omega)[h] = -\int_0^T \int_\Gamma (2[\![k]\!]\dot{u}\dot{p} + [\![c^2]\!]\nabla u \cdot \nabla p + [\![b]\!]\nabla \dot{u} \cdot \nabla p) h^\top n \, dx dt$$

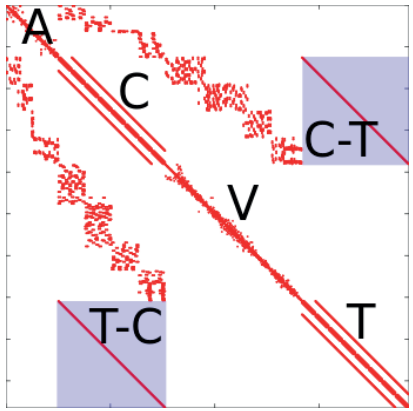


Uncertainty Quantification
The stochastic challenge

Peta-scale systems
The HPC challenge

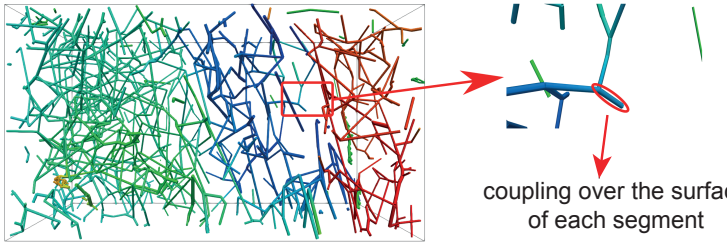
The analysis of PDEs is **fundamental** for developing efficient numerical schemes

Aspects of modern simulation technologies for PDEs

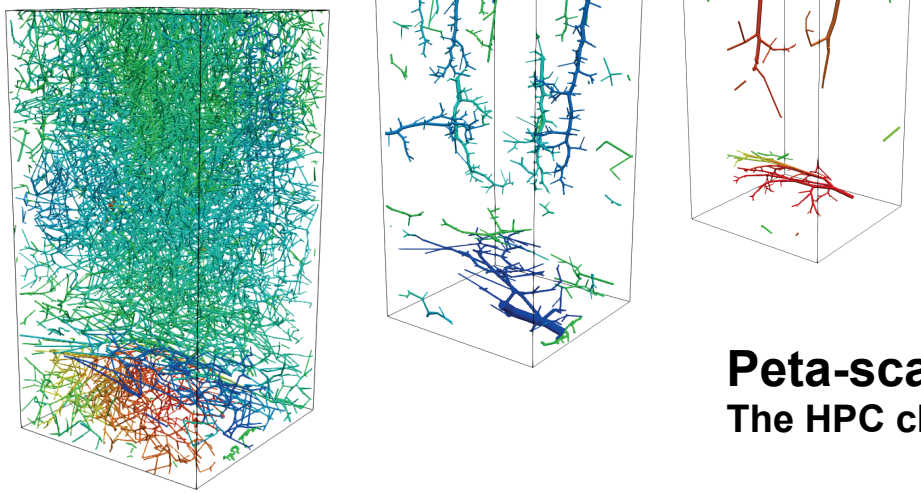


Mathematical modelling
The analysis challenge

capillary-tissue coupling



Numerical Analysis
The algorithmic challenge

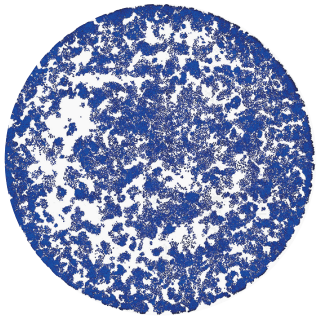


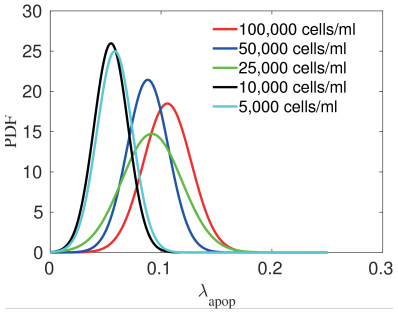
Uncertainty Quantification
The stochastic challenge

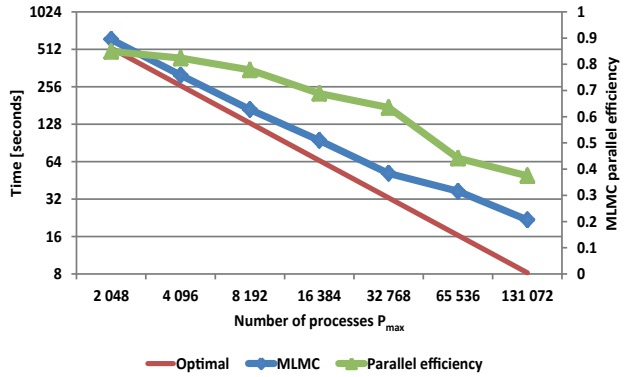
Peta-scale systems
The HPC challenge

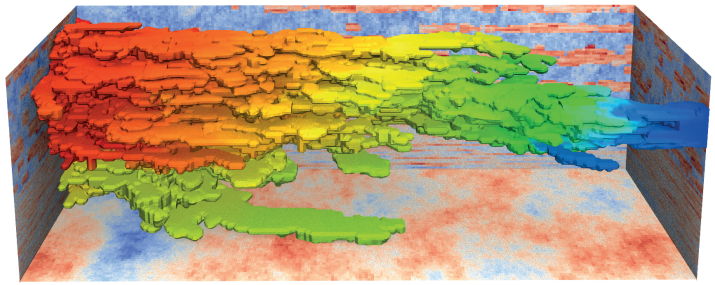
Multi-scale models are **essential** for predictive simulation of complex phenomena

Aspects of modern simulation technologies for PDEs

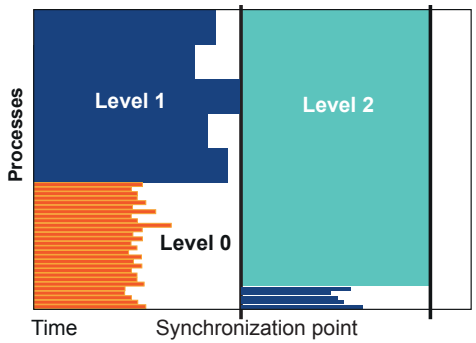








Mathematical modelling
The analysis challenge



Numerical Analysis
The algorithmic challenge

Uncertainty Quantification
The stochastic challenge

Peta-scale systems
The HPC challenge

Uncertainties increase **drastically** the computational complexity

Aspects of modern simulation technologies for PDEs

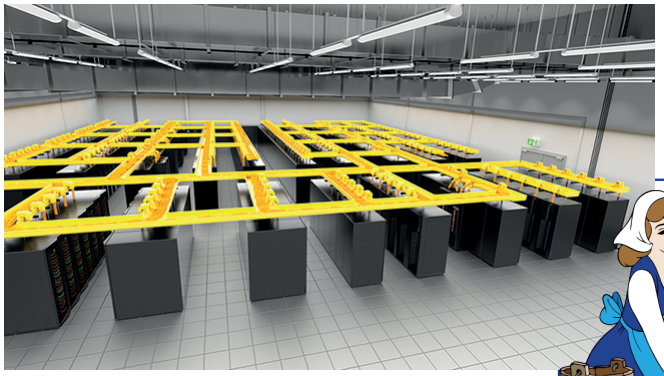
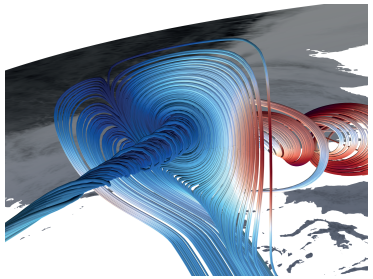
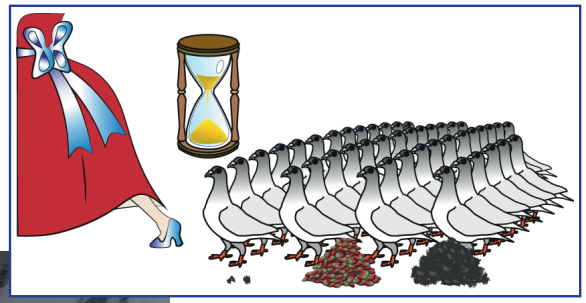
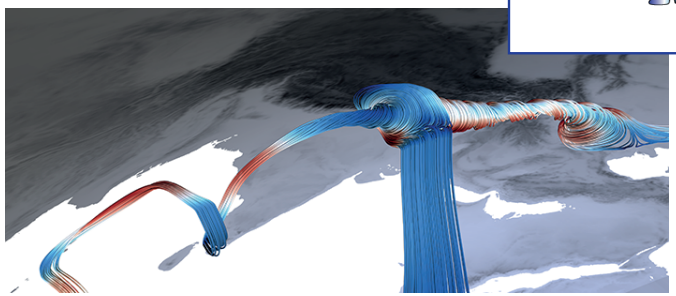


Foto: Leibniz Rechenzentrum







Mathematical modelling
The analysis challenge

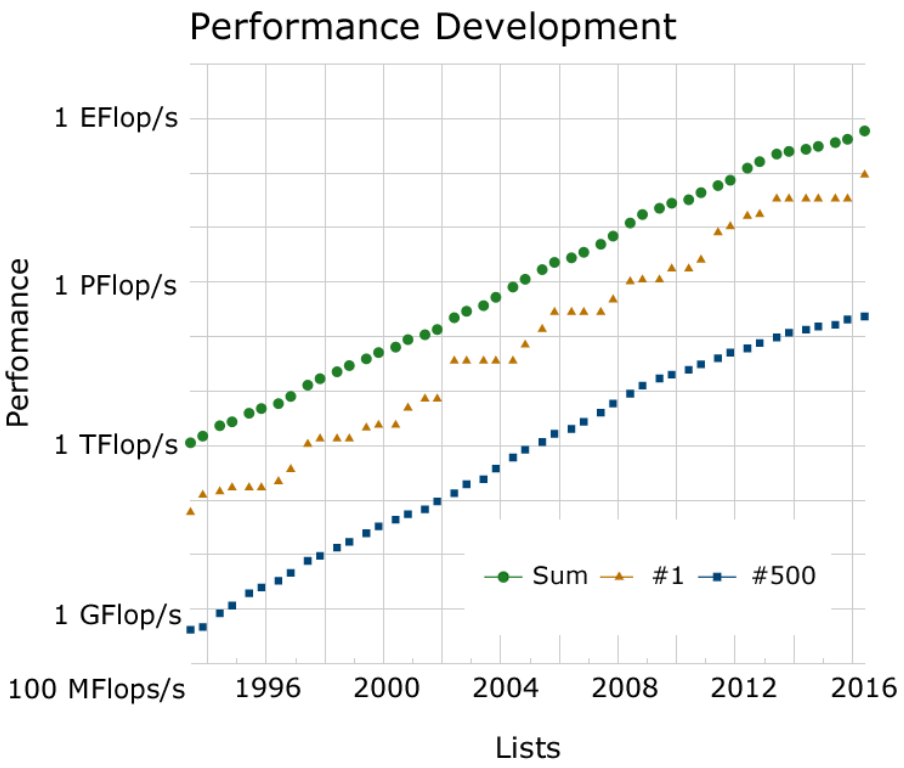
Numerical Analysis
The algorithmic challenge

Uncertainty Quantification
The stochastic challenge

Peta-scale systems
The HPC challenge

Scalable algorithms are **indispensible** for exploiting capabilities of HPC architectures

The computing power development



In the last two decades:

- Performance increase by $1\,000\,000$
- Memory increase by $10\,000$

Fastest supercomputer (Nov 2017):

The Sunway TaihuLight supercomputer (China), [Fu et al. 16]

- performance by a factor of ≈ 20, ≈ 500
- but only a factor of ≈ 3, ≈ 60 in memory

compared to
JUQUEEN (Germany), Hexagon (Norway)

Observations: The classical $\mathcal{O}(N^s)$ cost count metric is too simplistic.
Cost for communication and memory traffic **cannot** be ignored.

Some challenges for large scale FE

The geometry:
blending from reference to physical domain

The flow solver:
all-at-once MG for saddle point systems

The error control:
adaptivity beyond mesh refinement

Curved geometries in 3D: two-scale approach

Classical approach: element assembling – sparse matrix format – solve

Uniform refinement for **non-polyhedral** domains:

- **Cheap** and well-suited for on-the-fly
  but asymptotically wrong
- **Optimal** complexity and order
  but expensive

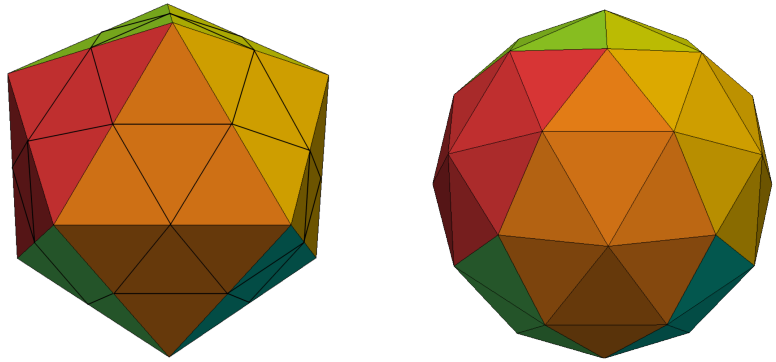
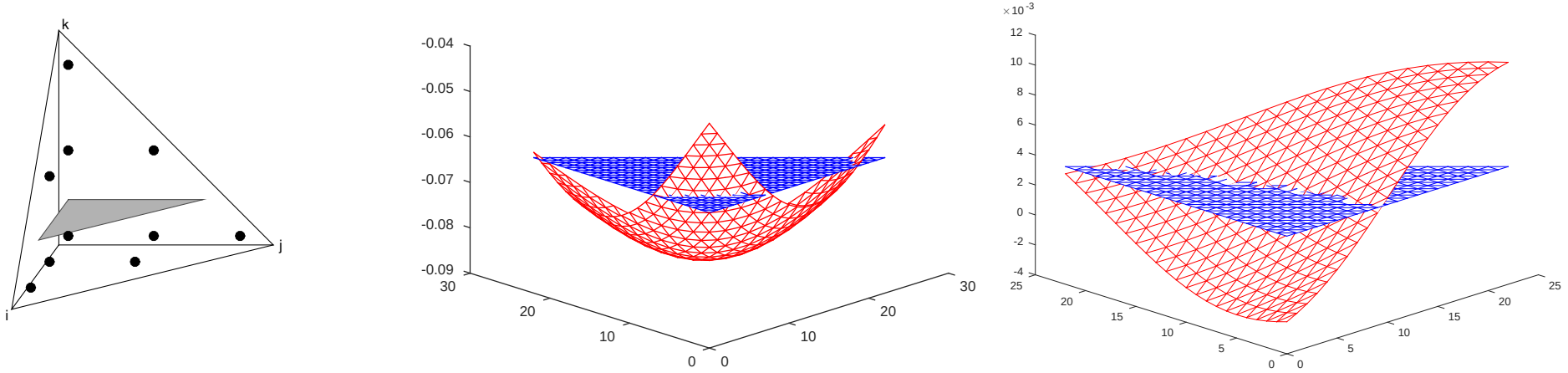


Illustration of two stencil entries as index functions over a 2D plane

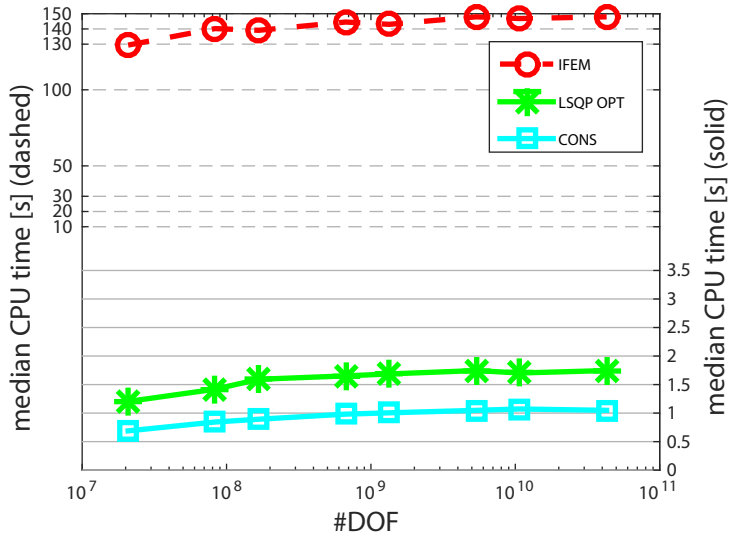


Observation: Stencil entries are smooth functions within each macro-element

Cost reduction versus accuracy loss (3D)

Idea: Replace the flop intense on-the-fly assembling by the evaluation of piecewise higher order **surrogate polynomials**

Observation: Drastic cost reduction compared to standard isoparametric FEM



Influence of the surrogate order and macro mesh-size on accuracy

To the right: Increase in the number of macro-elements from 60 to 30720

From top to bottom: Increase in the 3d refinement level from 2 to 6

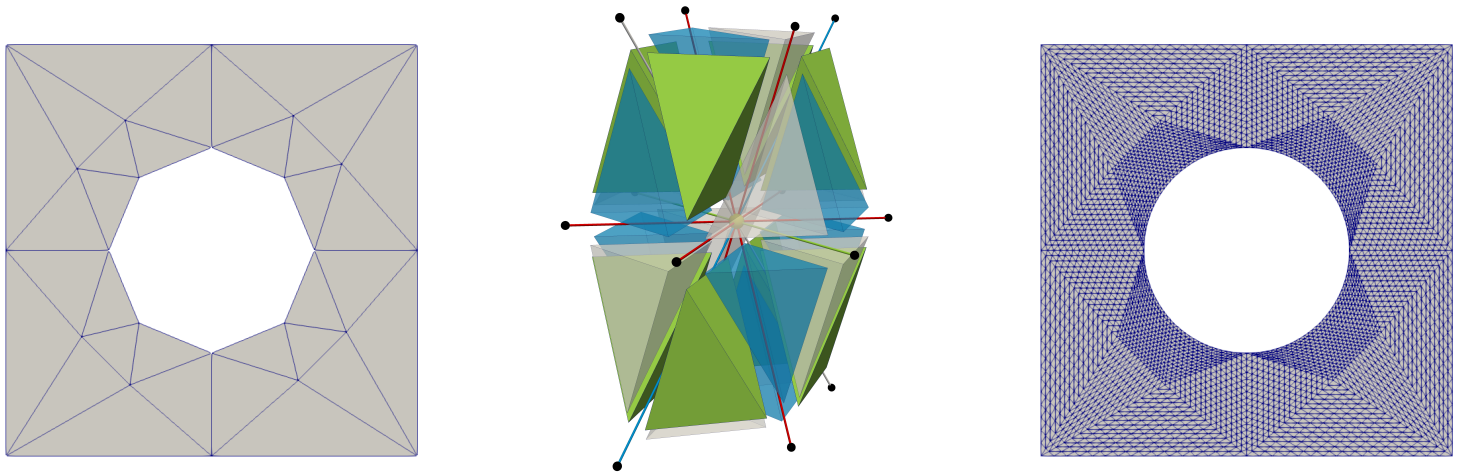
1.3e-03	3.6e-04	8.9e-05	2.3e-05
3.9e-04	9.4e-05	2.3e-05	5.7e-06
2.1e-04	2.5e-05	5.7e-06	1.4e-06
2.1e-04	*1.3e-05*	1.5e-06	3.6e-07
2.2e-04	*1.3e-05*	*5.6e-07*	9.0e-08
quadratic (upper) and cubic (lower)			
1.2e-03	3.6e-04	8.9e-05	2.3e-05
3.5e-04	9.4e-05	2.3e-05	5.7e-06
1.0e-04	2.4e-05	5.7e-06	1.4e-06
6.1e-05	6.2e-06	1.4e-06	3.6e-07
6.0e-05	*2.1e-06*	3.6e-07	8.9e-08

Abstract framework: transformation

- **Transfer** of the *physical* domain Ω by F onto the reference domain $\widehat{\Omega}$

$$s_{i,i+\epsilon_j} = \int_{\Omega} \nabla\phi_i \cdot K \nabla\phi_{i+\epsilon_j} = \int_{\widehat{\Omega}} \nabla\hat{\phi}_i \cdot \frac{DF \cdot K \cdot DF^T}{|\det DF|} \nabla\hat{\phi}_j$$

- **Exploit** a hybrid mesh structure, i.e. unstructured initial mesh and uniform refinement



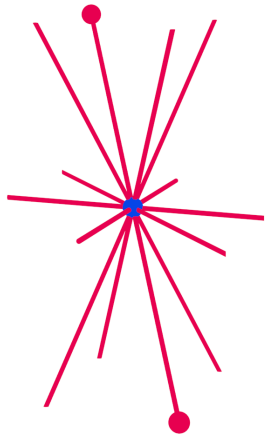
unstructured coarse mesh (2D) – structured stencil (3D) – transformed uniformly refined mesh (2D)

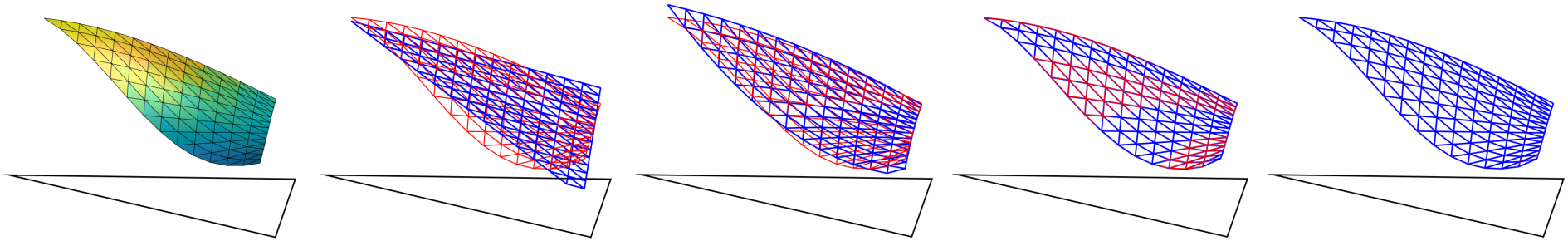
- **Replace** the stiffness matrix entries per macro-element by a surrogate polynomial

Abstract framework: approximation

Convolution: $s_{i,i+\epsilon_j} = \int \nabla\phi_0(x)\hat{K}(x-p_i)\nabla\phi_j(x)\,dx$

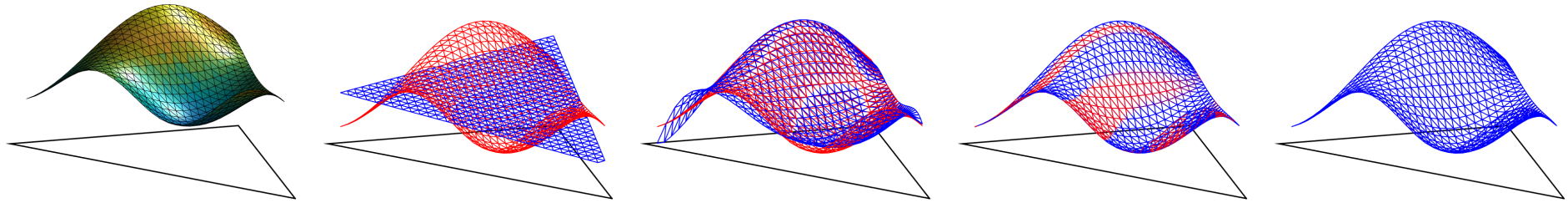
Observation: If transformed coefficient $\hat{K}$ is a tensor with polynomial entries then the stencil entry at node p in direction $\pm\epsilon_j$ is a polynomial of the same degree.





Coefficient (left) and $q = 1, 2, 4$, and 7 (from left to right)

Question: How to choose the polynomial degree and number of surrogate polynomials?



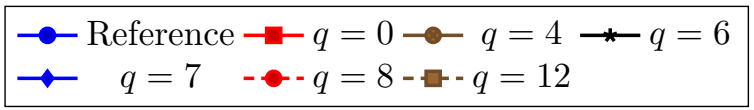
Coefficient (left) and $q = 1, 4, 8$, and 12 (from left to right)

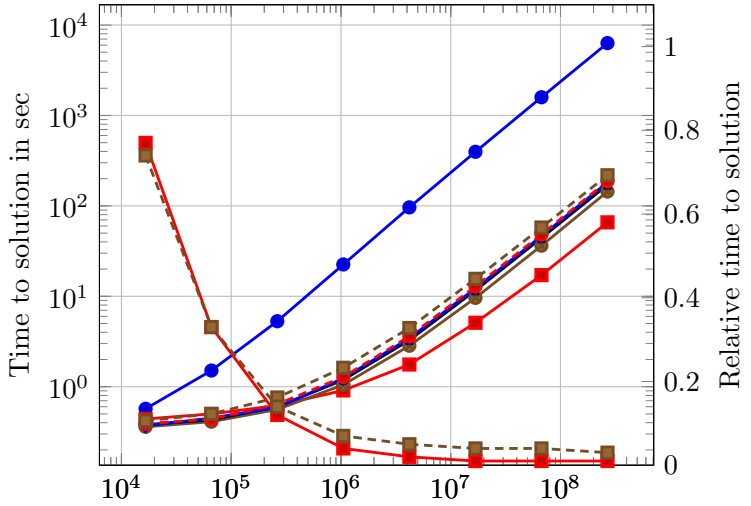
Numerical results

- Setup (left) and **run-time comparison** (right)

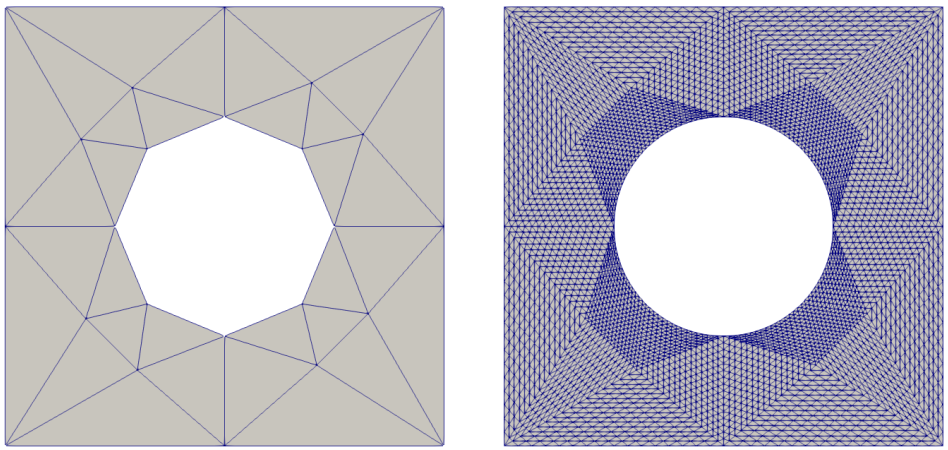
$$-\mathrm{div}\,(K \cdot \nabla u) = f \quad \text{in } \Omega + \text{ BC}$$

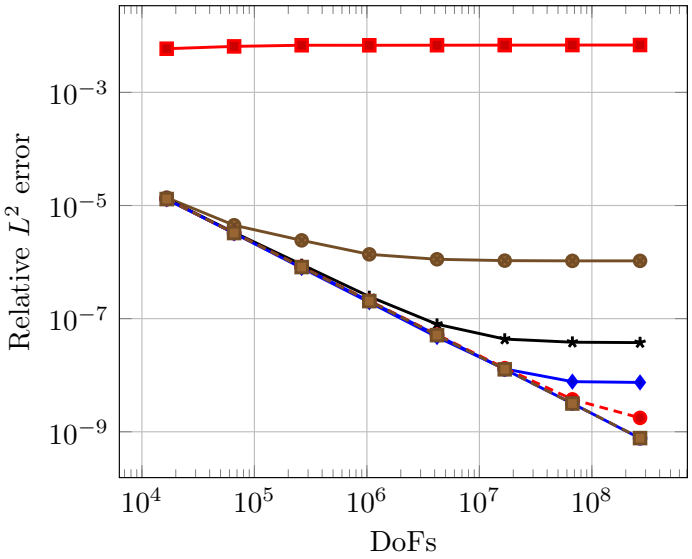
$$K = \begin{pmatrix} 3x^2 + 2y^2 + 1 & -x^2 - y^2 \\ -x^2 - y^2 & 4x^2 + 5y^2 + 1 \end{pmatrix}$$





- **Accuracy comparison** for fixed H with respect to q





Convergence rates with respect to H

Theorem: The H^1-**discretization** error is given by

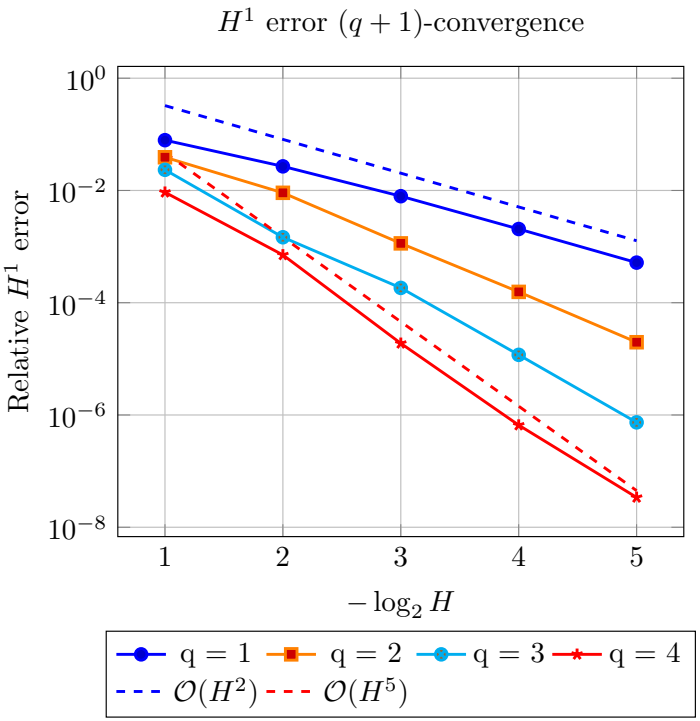
$$\|\nabla(u - \tilde{u}_h)\| = \mathcal{O}(h^p) + \mathcal{O}(H^{q+1})$$

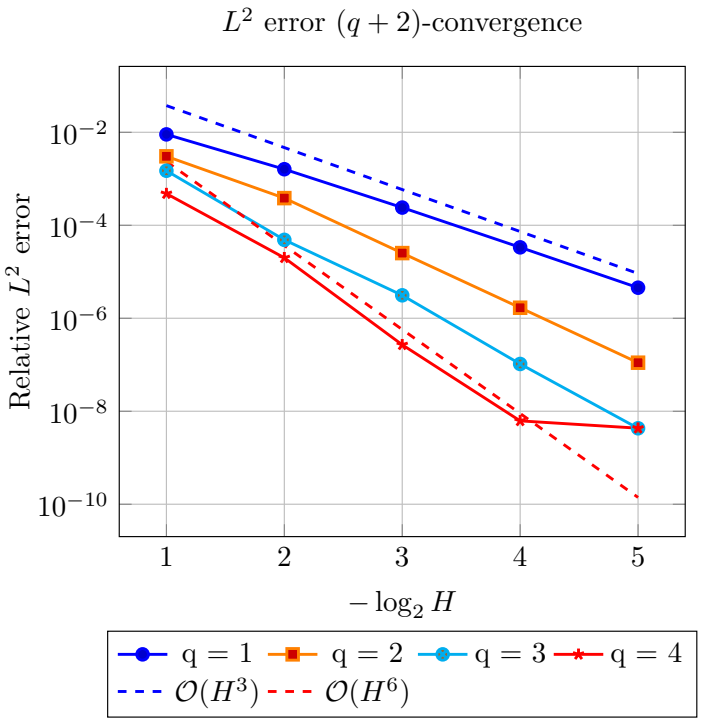
The L^2-**discretization** error is given by

$$\|u - \tilde{u}_h\| = \mathcal{O}(h^{p+1}) + \mathcal{O}(H^{q+2})$$

$p :=$ finite element order, $q :=$ surrogate polynomial degree

$h :=$ finite element basis support diameter, $H :=$ surrogate polynomial support diameter





Abstract framework: theory and control

Question: How to choose the required polynomial degree q?

Idea: Increase q **adaptively** within an iterative multigrid solver

Basic algorithm

1. Perform V-cycles with a fixed polynomial degree q until stopping Criteria 1 is satisfied
2. Increase polynomial degree $q \leftarrow q + 1$ and perform an additional V-cycle
   stop if the update satisfies stopping Criteria 2; otherwise go to step 1

Remarks on the selection of the stopping criteria:

Criteria 1: residual, estimator for algebraic error, estimator for total error
Criteria 2: as above or difference to the updated solution with respect to $q + 1$

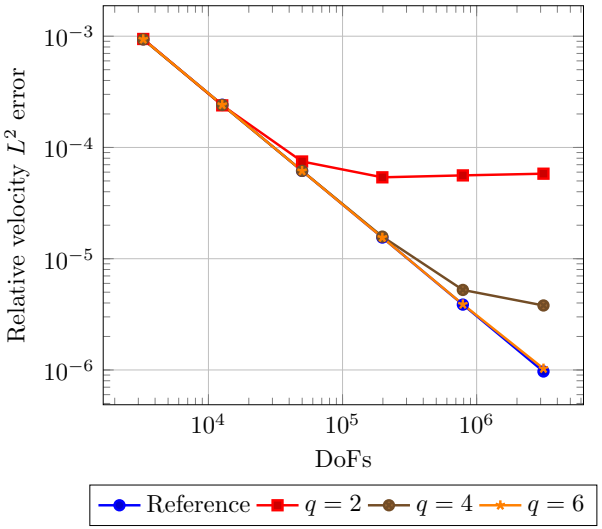
	standard FE		adaptive surrogate FE			
mesh level	L^2 err.	time [s]	L^2 err.	final q	time [s]	ratio
5	8.69e-06	1.09	8.83e-06	5	0.41	38 %
6	2.18e-06	2.89	2.55e-06	5	0.65	22 %
7	5.47e-07	9.48	5.85e-07	6	2.09	22 %
8	1.37e-07	32.93	1.83e-07	6	4.92	15 %
9	3.42e-08	136.78	3.00e-08	7	18.19	13 %

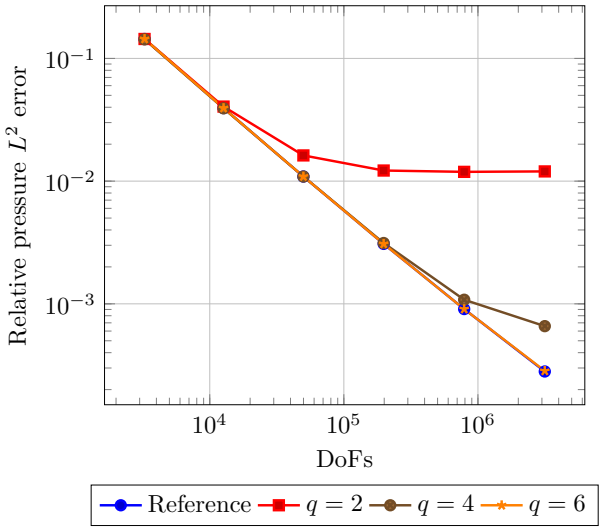
Surrogate polynomials for more general settings

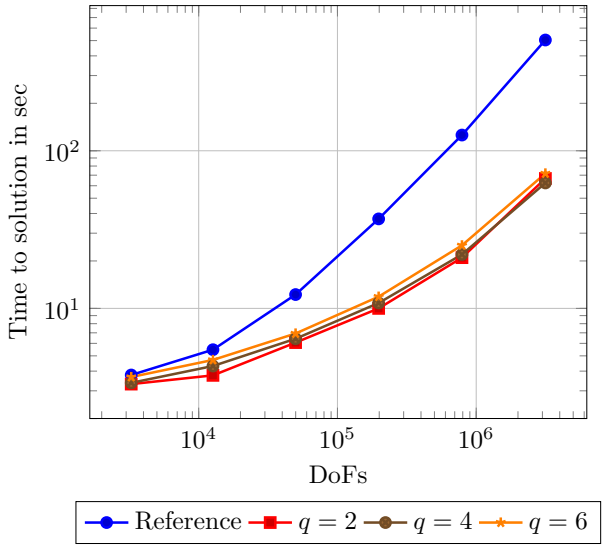
- The **Darcy** case for $P2$ finite elements

L	Err. ref.	eoc	tts [s]	Err. $q = 4$	Err. $q = 6$	Err. $q = 8$	tts [s]
2	3.67E-05	-	0.00	3.63E-05	3.67E-05	3.67E-05	0.00
3	2.94E-06	3.64	0.01	2.96E-06	2.95E-06	2.94E-06	0.00
4	2.09E-07	3.81	0.09	6.28E-07	2.10E-07	2.09E-07	0.02
5	1.40E-08	3.91	0.82	4.85E-07	1.41E-08	1.40E-08	0.10
6	9.01E-10	3.95	6.91	4.68E-07	1.16E-09	9.01E-10	0.61

- The **Stokes** case with stabilized P1-P1 elements

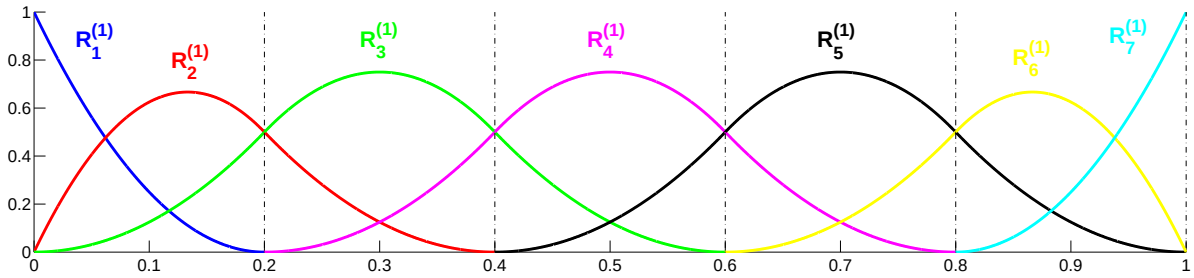






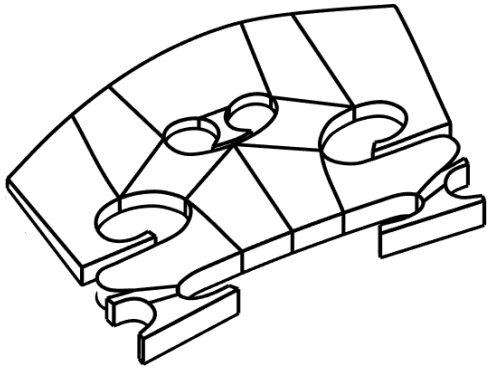
Patch-wise isogeometric elements

- **Use** standard isogeometric elements on each patch



- B-splines S_h^p
- maximal regularity
- tensorial geometries

- **Impose** weak continuity conditions at interfaces



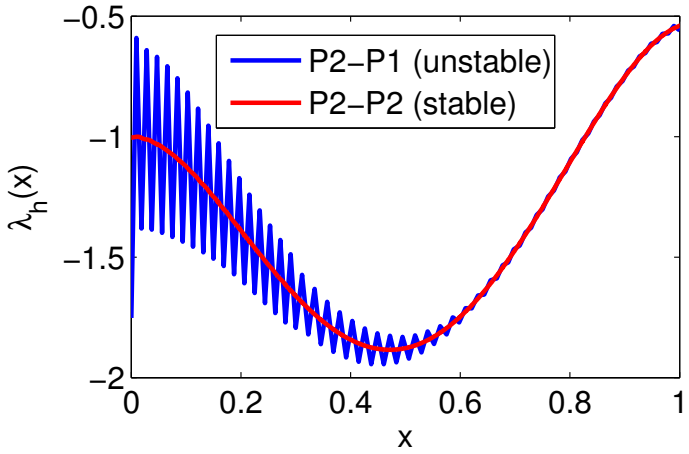
- Lagrange multipliers M_h
- reference domain
- physical domain

- **Saddle point** formulation in displacement and surface traction

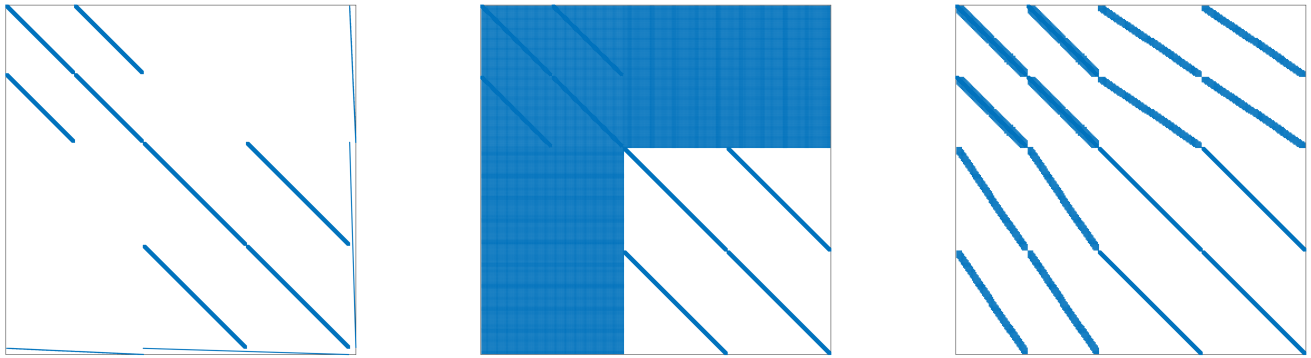
Elementary references: [de Boor 01], [Schumaker 07], [Cottrell, Hughes, Bazilevs 05], [Höllig 03]

How to define the Lagrange multiplier space?

- **Theoretical aspect:** reproduction property of order $p-1$ and inf-sup stability



- **Computational aspect:** local support and biorthogonality

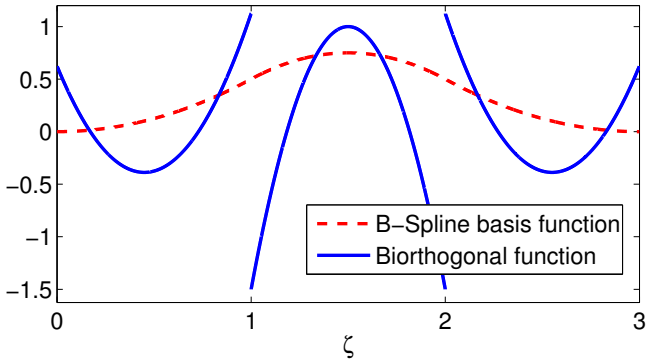


saddle point (left), standard LM (middle) and biorthogonal (right)

allow for local static condensation and symmetric and positive definite system

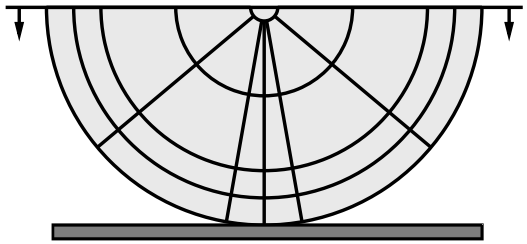
Biorthogonal basis functions

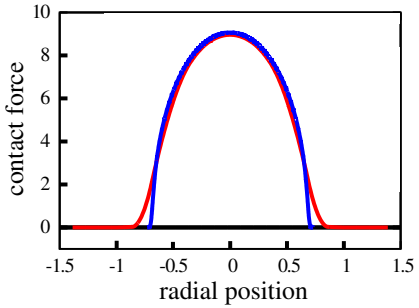
Straightforward computation by a local inversion
Low order approximation properties, but optimal for contact problems:



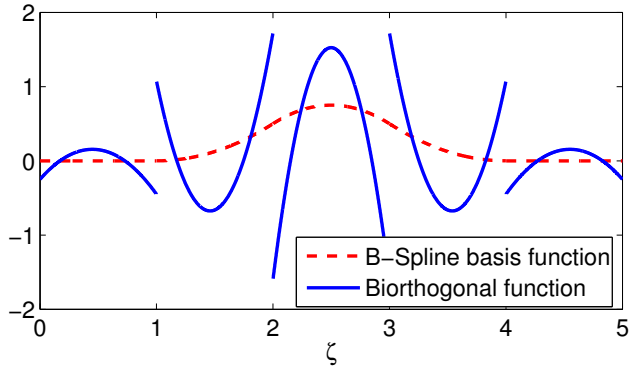
Good results for **contact problems** in nonlinear elasticity

[Seitz, Farah, Kremheller, W, Popp, Wall 16]



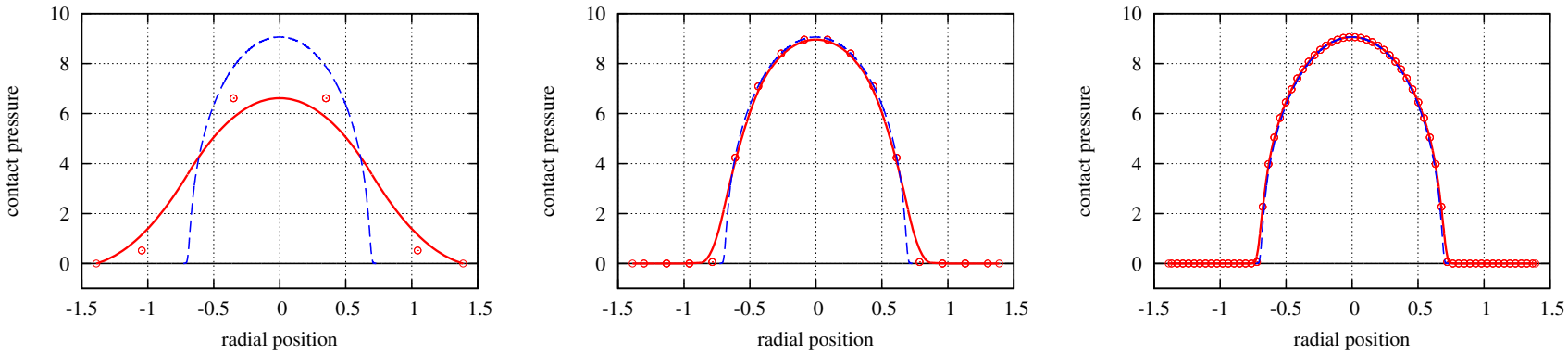


Alternative computation with enlarged support
Optimal approximation properties, suitable for domain decomposition [Oswald, W 02]

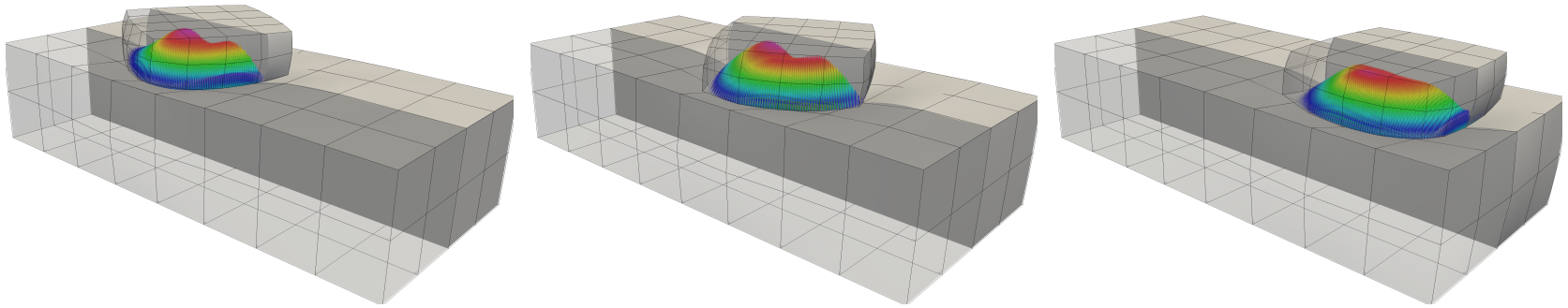


Biorthogonal IGA for contact mechanics

- **Mesh study** for classical Hertz example (qudaratic NURBS)



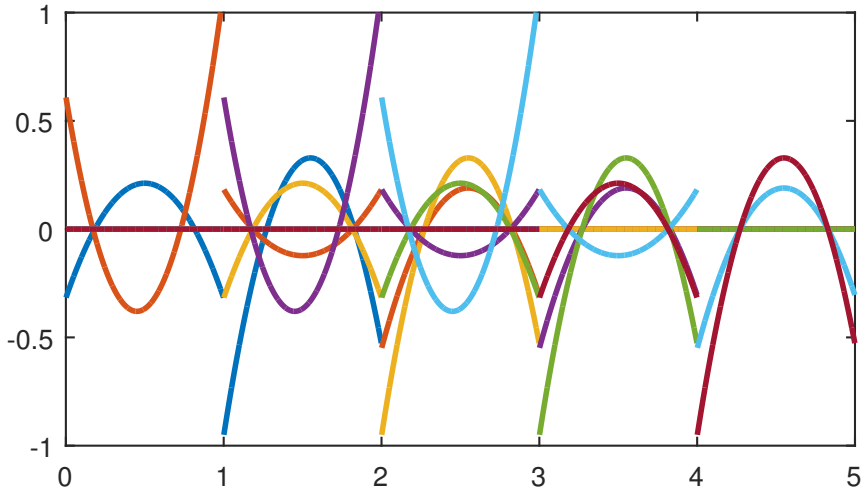
- **Rotating ironing** with Coulomb friction and neo-Hookean hyperelastic law



Optimal convergence rates due to reduced regularity of the solution resulting from the quasi-variational inequality (reproduction property is limited to P_0 for the LM space)

A locally constructed biorthogonal basis

- **Define** the coefficients **locally** by the inverse of the element mass matrix
- **Glue** the basis functions **globally** together such that supp $N_i =$ supp $\widetilde{\psi}_i$
- **Idea:** enrich the space by orthogonal functions of the same order, $\int N_i z_j \, dx = 0$
- **Follow** the FEM case [Oswald, W 02] and enlarge the support



z_i for $p = 2$

$$\boldsymbol{\psi}_i = \widetilde{\boldsymbol{\psi}}_i + \boldsymbol{\alpha}_{ij}\mathbf{z}_j$$

How to define α_{ij} such that local support and $p-1$ reproduction property hold?

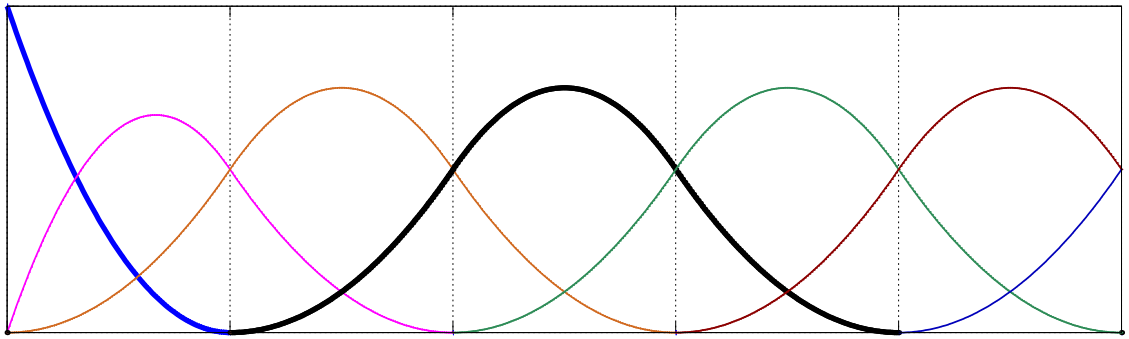
The choice of the coefficient

- The **locality** of the support: Define left center element e_i

$\alpha_{ij} = 0$ if $N_j = 0$ on e_i

results in a support of:

$2p + 1$ elements



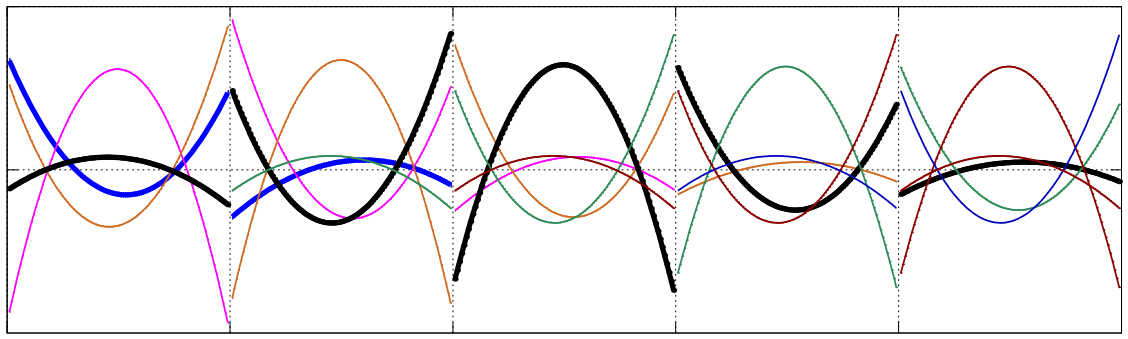
center element

- The **reproduction** property: Solve a local system for α_{ij}, $i \in I_j$, $\#I_j = p + 1$

$$\sum_{i \in I_j} (p_l, N_i)\alpha_{ij} = (p_l, \phi_j)$$

p_l, $l = 1, \ldots p + 1$ basis of P_p

ϕ_j suitable basis of product space



dual basis with extended support

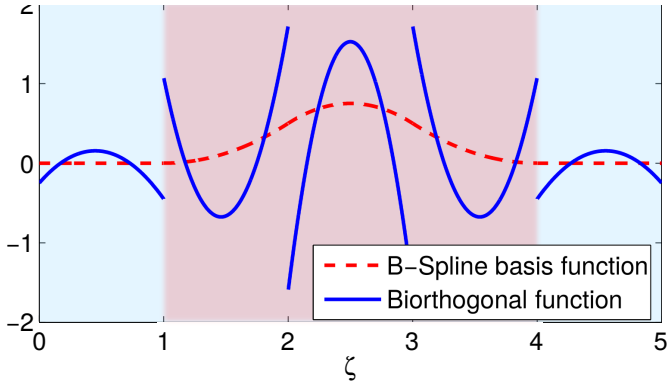
Lemma: The quasi-interpolant $Qf := \sum_i (f, N_i)\psi_i$ is then **invariant** for P_p

Biorthogonal basis functions

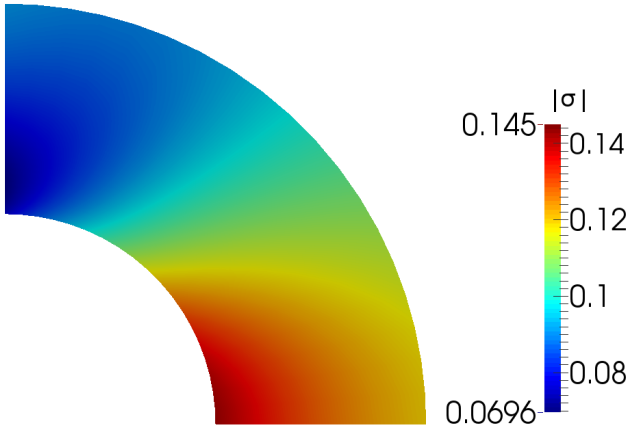
- **primal** support: $p+1$ elements
- **dual** support: $2p+1$ elements

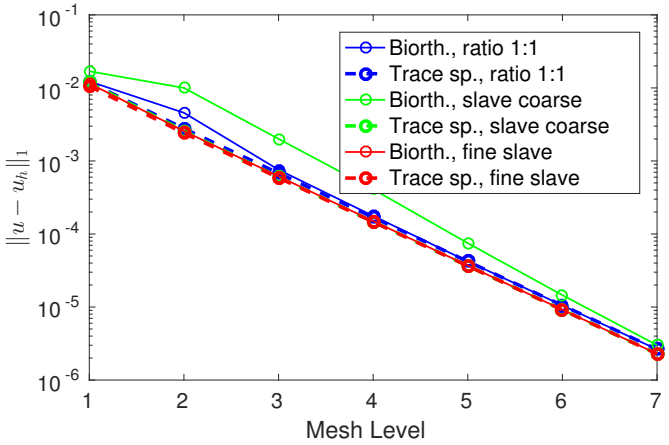
[Wunderlich et al, W 18]

$p = 2$:

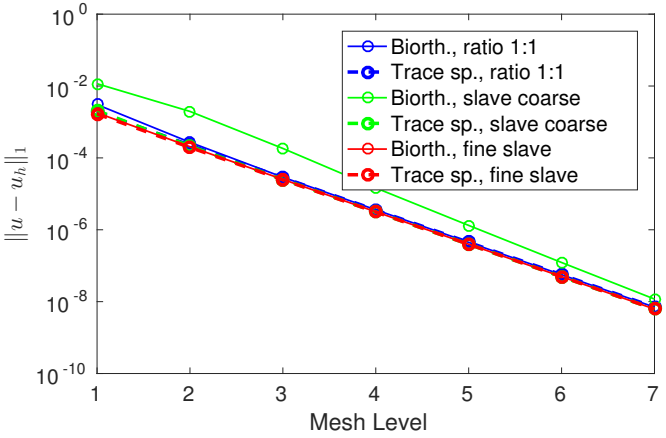


Annulus domain: comparing to trace spaces, different mesh-ratio





$p = 2$



$p = 3$

Optimal convergence rate and local static condensation by biorthogonal basis

Application to nonlinear elasticity

Neo-Hookian material:

$$\text{Div}\,\mathbf{FS} + \mathbf{b} = \mathbf{0} \quad \text{in } \Omega$$

$$\mathbf{u} = \mathbf{0} \quad \text{on } \Gamma_D, \mathbf{FSN} = \mathbf{t} \quad \text{on } \boldsymbol{\Gamma_N}$$

$\mathbf{S} = \mathbf{2}\partial\boldsymbol{\Psi}/\partial\mathbf{C}$ second Piola–Kirchhoff stress

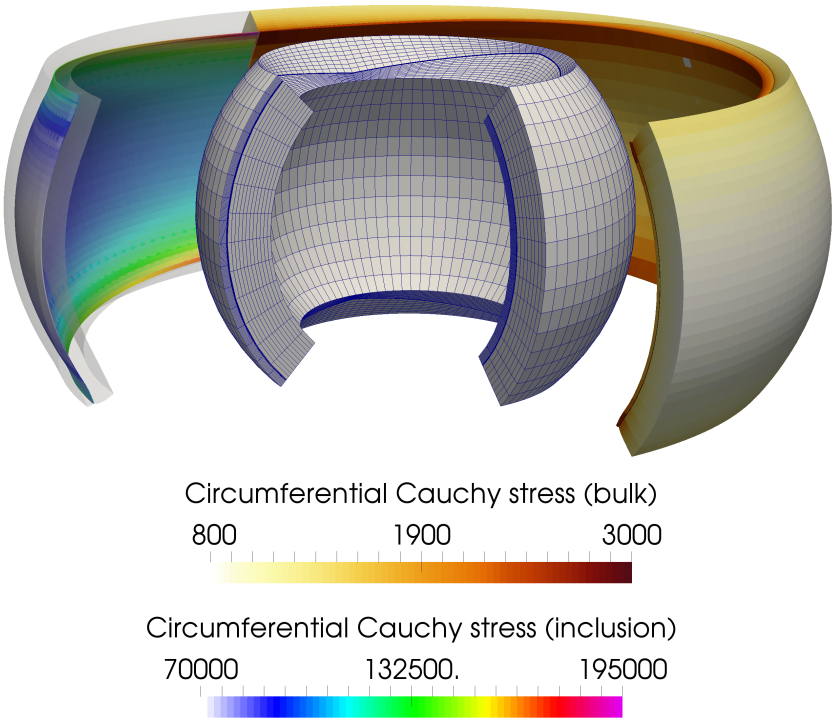
$\mathbf{C} = \mathbf{F}^\top\mathbf{F}$ right Cauchy–Green tensor

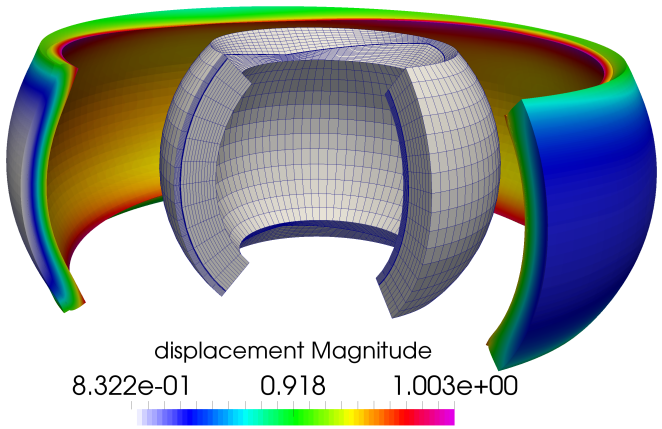
$\mathbf{F}$ deformation gradient

$\boldsymbol{\Psi}$ strain energy function

$\boldsymbol{\Psi}(\mathbf{C}) = c(\text{tr}\,\mathbf{C} - 3) + \frac{c}{\beta}((\det\mathbf{C})^{-\beta} - 1)$

c, β material constants





Domain: Spherical shell with a $45°$-segment removed on the top and bottom

Inclusion: stiffer material on a thin elliptical crossection

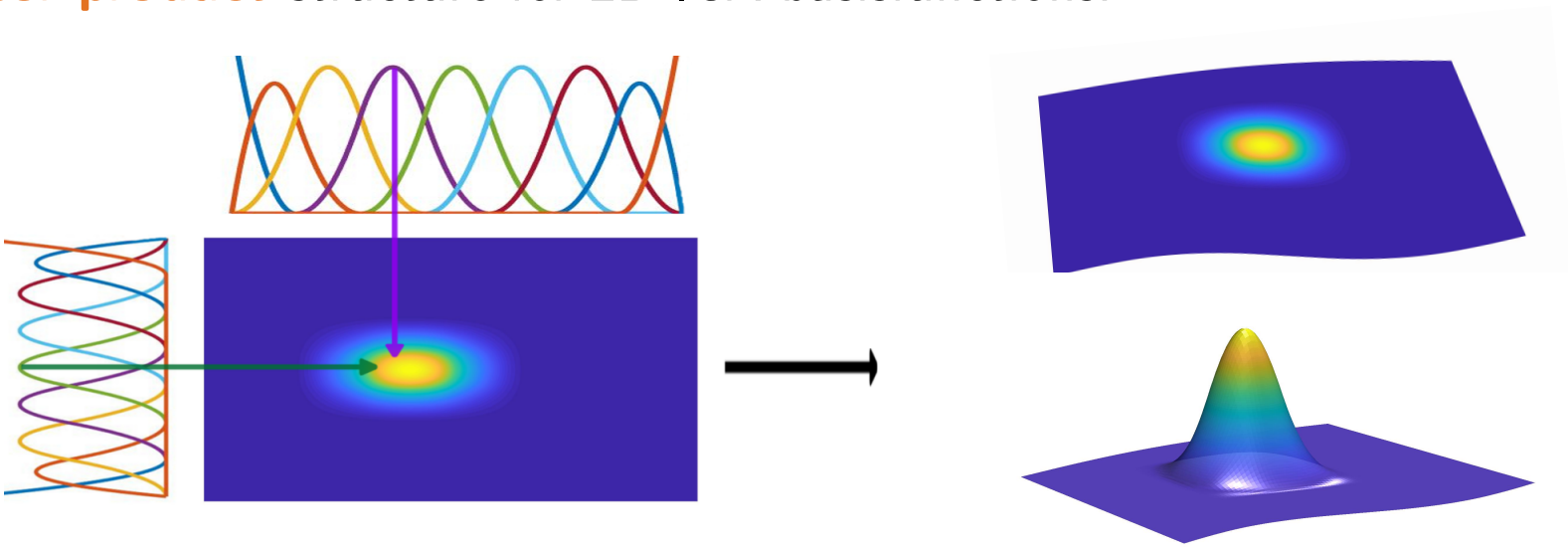
Discretization: quadratic NURBS

Control points: 104016

Polynomial stencil approximation

Classical assembling for IGA is quite expensive

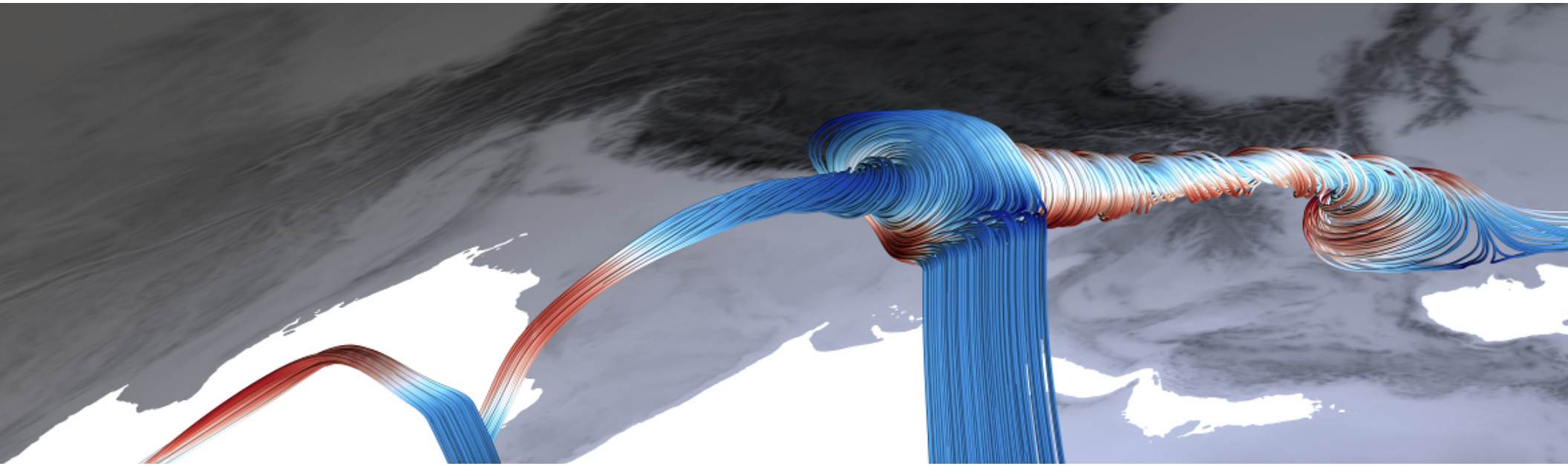
- **Tensor product** structure for 2D IGA-basisfunctions:



- **Stiffness matrix** entry $K_{i,i+1}$:

Drastic cost reduction in assembling possible by surrogate matrix

Surrogate FE operators in large scale simulation



partly joint work with W. Zulehner (2018)

All-at-once multigrid solver

Abstract saddle point system: $\begin{bmatrix} A & B^\top \\ B & -C \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}$

Different solver strategies:

- preconditioned Krylov space solver (e.g. minres) for indefinite system
- preconditioned Krylov space solver (e.g. cg) for positive definite Schur complement
- **all-at-once** multigrid for indefinite system

Different smoother strategies:

- Braess–Sarazin type [Braess et al 97]
  global saddle point structure
- Vanka type [Vanka 86], [Manservisi 06]
  local saddle point structure
- **Uzawa** type [Gaspar et al. 14], [Zulehner 00-03]
  smaller flop and communication count

$$\begin{bmatrix} \hat{A} & 0 \\ B & -\hat{S} \end{bmatrix}$$

Convergence result

Smoothing property [Drzisga et al., W 18]
Assume that A is symmetric and positive definite. Let $\hat{A}$ and $\hat{S}$ be symmetric and positive definite matrices satisfying

$$\hat{A} \geq A \quad \text{and} \quad \hat{S} \geq S := C + BA^{-1}B^{\top},$$

then the following smoothing property holds for a Uzawa type iteration:

$$\|\mathcal{A}\mathcal{M}^{\nu}\|_{\mathcal{L}\times\mathcal{L}} \leq \sqrt{2}\,\eta(\nu-1)\,\|\mathcal{D}_d\|_{\mathcal{L}\times\mathcal{L}}, \quad \nu \text{ number of smoothing steps}$$

with $\mathcal{D}_d = \begin{bmatrix} \hat{A} & 0 \\ 0 & \hat{S} \end{bmatrix}$, $\mathcal{M} := \mathrm{Id} - \begin{bmatrix} \hat{A} & 0 \\ B & -\hat{S} \end{bmatrix}^{-1} \mathcal{A}$ and $\eta(\nu) = \frac{1}{2^{\nu}} \binom{\nu}{\lfloor(\nu+1)/2\rfloor}$.

Proof: Based on abstract framework of Reusken [91].

Theorem: Level independent W-cycle convergence results are then guaranteed.

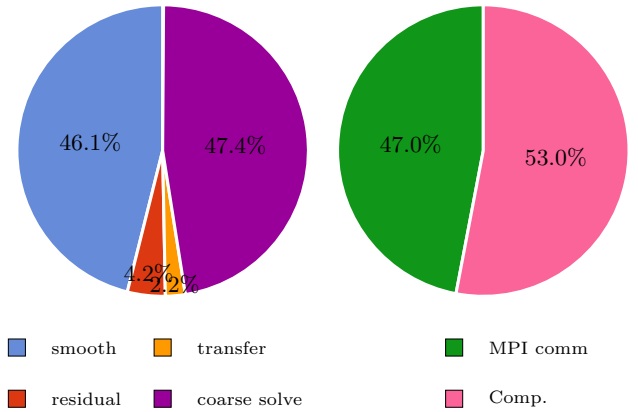
Remark: The theory can be extended to a variable V-cycle but not to the V-cycle. Numerical results show that the theory is sharp.

Parallel efficiency on JUQUEEN

Observation: Parallel efficiency is significantly reduced for huge systems

Step 1: Replace non-scalable Krylov coarse mesh solver by a PETSc solver

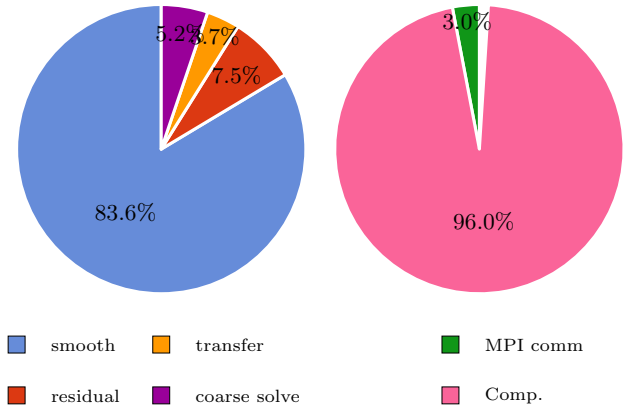
- Setup phase: Save matrix in standard CRS format
- MINRES-iteration with block preconditioner
- velocity: AMG preconditioned CG-iteration
- pressure: lumped mass-matrix
- GAMG V(1,0), Chebyshev, 5 lev., threshold 0.01



Step 2: master-slave agglomeration on coarse level

np	DOF	red.	T[s]	coarse	par. eff
30	$8.3 \cdot 10^7$	1	16.284	0.043	1.00
120	$3.3 \cdot 10^8$	1	16.426	0.050	0.99
960	$2.6 \cdot 10^9$	1	17.084	0.171	0.95
7680	$2.4 \cdot 10^{10}$	1	17.310	0.382	0.94
61440	$1.7 \cdot 10^{11}$	8	17.704	0.877	0.92

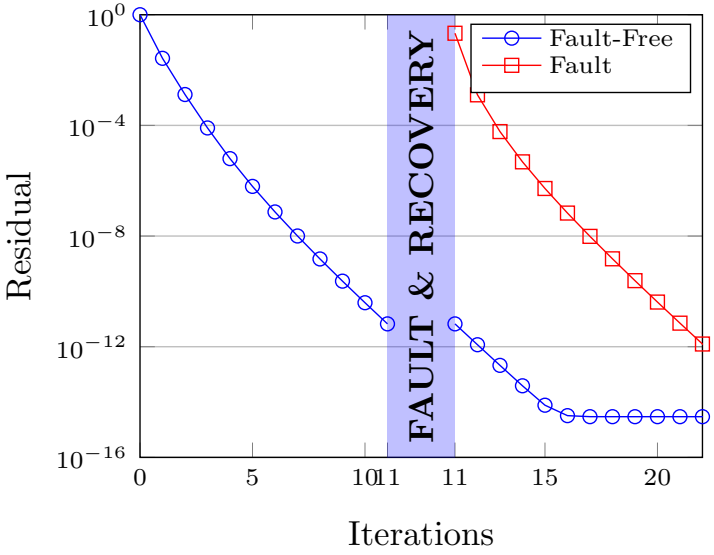
np: 61440



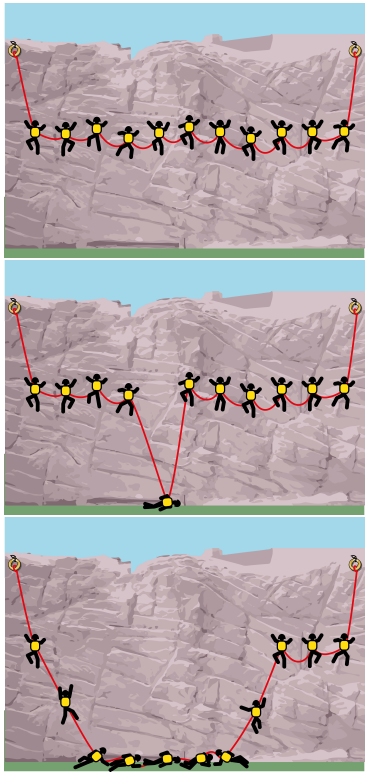
results in a **parallel efficiency** of more than 90%.

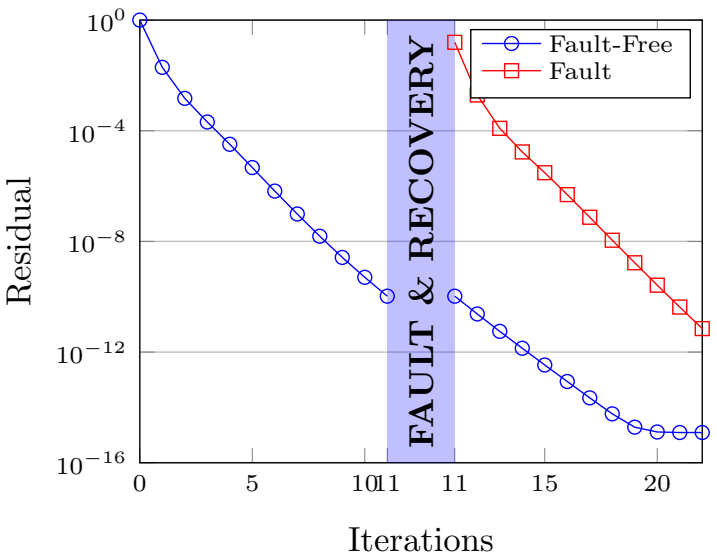
Adaptive error control for MG solvers

- **No recovery at all:**



Laplace





Stokes

- **Algorithmic recovery strategy:**
  - **Freeze** the data on the adjoint lower primitives (Dirichlet IC)
  - **Replace** the faulty processor by several ones (over balancing)
  - **Control** the catch-up progress (hierarchical residual representation)

Dirichlet-Dirichlet recovery strategy

Dirichlet boundary condition on **healthy** and on **faulty** domain

1: Solve $Au = f$ by multigrid cycles.
2: **if** Fault has occurred **then**
3: **STOP** solving.
4: Recover **Dirichlet** boundary data u_{Γ_F} from row 4
5: Initialize inner values u_F with zero
6: **In parallel approximate Dirichlet problem on subdomains:**
7: Use $\mathbf{n_F}$ MG cycles accelerated by **superman** η_s to approximate row 5:
8: $A_{FF}u_F = f_F - A_{F\Gamma_F}u_{\Gamma_F}$
9: Use $\mathbf{n_I}$ MG cycles to approximate row 1
10: $A_{II}u_I = f_I - A_{I\Gamma_I}u_{\Gamma_I}$
11: **RETURN** to row 1 with new values u_I in Ω_I and u_F in Ω_F.
12: **end if**

$$\begin{pmatrix} A_{II} & A_{I\Gamma_I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Id} & \mathbf{-Id} & \mathbf{0} & \mathbf{0} \\ A_{\Gamma I} & \mathbf{0} & A_{\Gamma\Gamma} & \mathbf{0} & A_{\Gamma F} \\ \mathbf{0} & \mathbf{0} & \mathbf{-Id} & \mathbf{Id} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & A_{F\Gamma_F} & A_{FF} \end{pmatrix} \begin{pmatrix} u_I \\ u_{\Gamma_I} \\ u_\Gamma \\ u_{\Gamma_F} \\ u_F \end{pmatrix} = \begin{pmatrix} f_I \\ 0 \\ f_\Gamma \\ 0 \\ f_F \end{pmatrix}$$

Question: How to select $\mathbf{n_F}$ and $\mathbf{n_I}$?

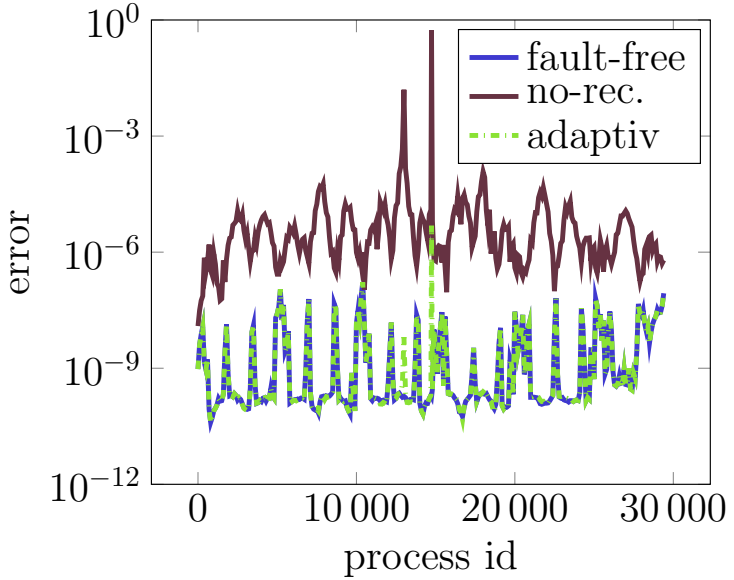
Adaptive tearing and interconnecting

- Error **control** (left) and **local** distribution (right)

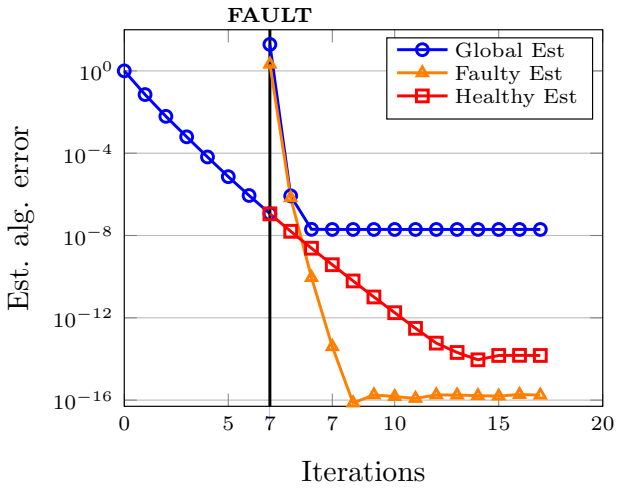
Hierarchical weighted residual: [Rüde 93]

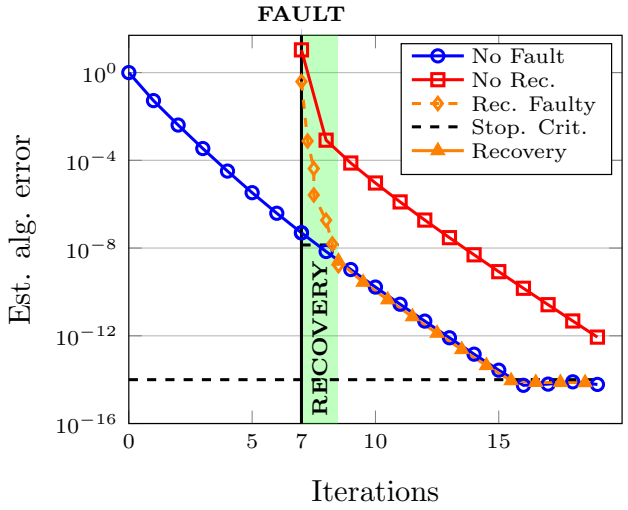
$$\eta := \| \sum_{l=0}^{L} I_l^L D_l^{-1} I_L^l r_L \|$$

Decouple over the interface and use in the faulty domain only the **inner** residual



- **No** recoupling (left) and **adaptive** recoupling for a superman factor of four (right)





The mantle convection model

The **physical model** consists of conservation of **momentum, mass and energy**

$$
\begin{aligned}
-\operatorname{div}\boldsymbol{\sigma} &= \rho\mathbf{g} \\
\operatorname{div}(\rho\mathbf{u}) &= 0 \\
\partial_t(\rho e) + \operatorname{div}(\rho e\mathbf{u}) &= -\operatorname{div}\mathbf{q} + \rho H + \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}
\end{aligned}
$$

Key quantities: *velocity* $\mathbf{u}$, *temperature* T, *pressure* p, and the *mantle viscosity* μ. The *density* ρ is given by the **mineralogy** via an equation of state:

$$\rho = \rho(p, T)$$

The **rheology** of the mantle is an active field:

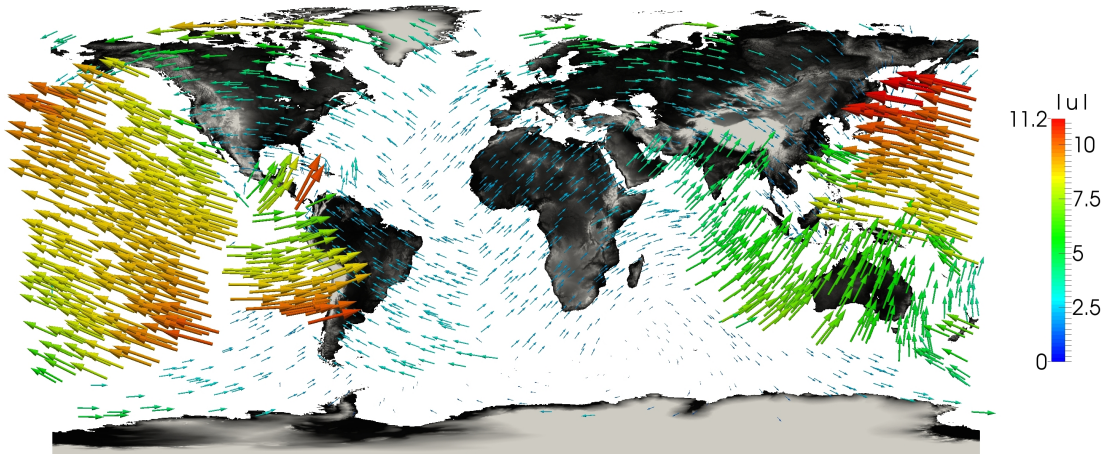
$$\boldsymbol{\sigma} = 2\mu(\dot{\boldsymbol{\varepsilon}} - \tfrac{1}{3}\operatorname{tr}\dot{\boldsymbol{\varepsilon}}\cdot\mathbf{I}) - p\mathbf{I}, \quad \text{with } \mu = \mu(r, T, \dot{\boldsymbol{\varepsilon}})$$

Notation:

- $\boldsymbol{\sigma}$ stress tensor
- ρ density
- $\mathbf{g}$ gravitational acceleration
- $\mathbf{u}$ velocity
- e internal energy
- $\mathbf{q}$ heat flux per unit area
- H volumetric radiogenic heat production rate
- $\dot{\boldsymbol{\varepsilon}}$ rate of strain tensor

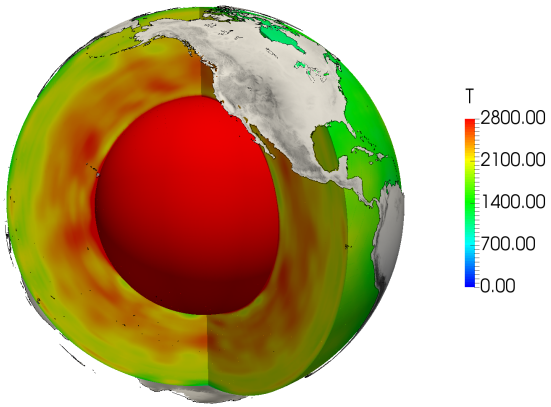
Data from measurements

- **Outer velocity boundary conditions:** Plate tectonic reconstruction

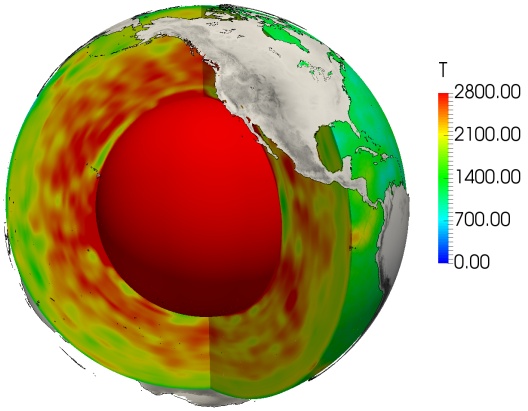


Williams, Müller, Landgrebe, Whittaker:
GSA Today, 2012

- **Temperature data:** Representation of seismic data by spherical harmonics:



Grand, van der Hilst, and Widiyantoro:
GSA Today, 1997

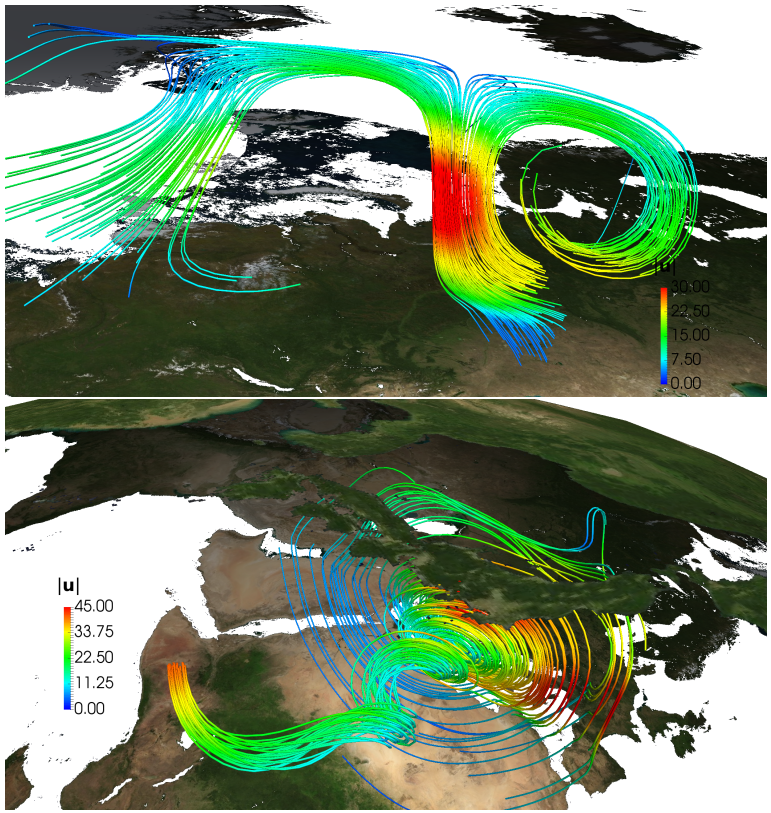


Simmons, Myers, Johannesson, Matzel, and Grand:
Geophysical Research Letters, 2015

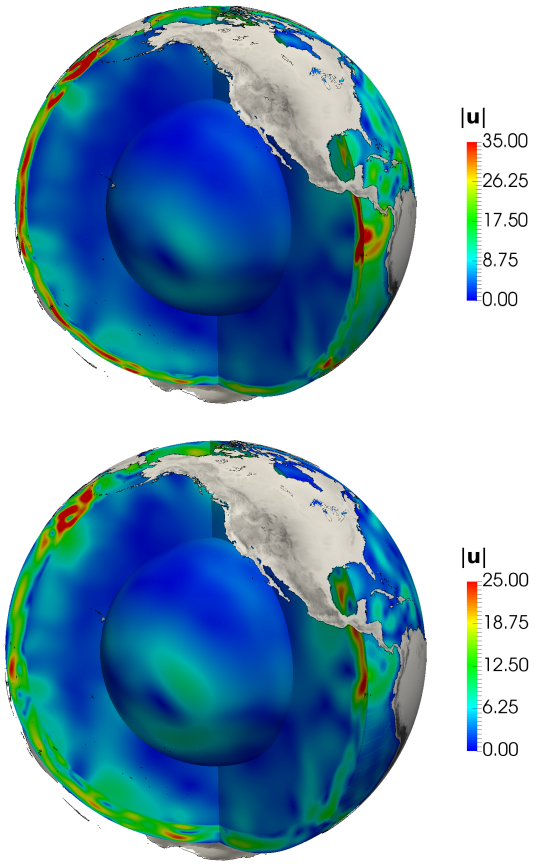
Temperature and depth dependent rheology

Viscosity model according to [Davies et al. 2012]: $d_A := 410,660$

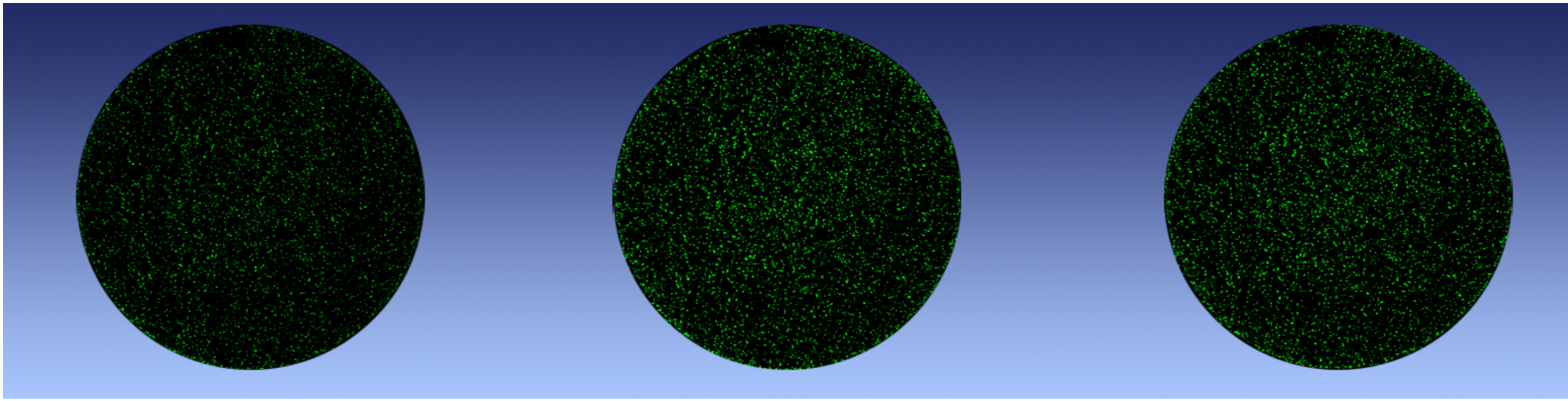
$$\mu(r,T) = \exp(4.61\,\tfrac{1-r}{1-r_{\text{inner}}} - 2.99\,T) \begin{cases} \frac{1}{10}\,d_{\text{a}}^3 & \text{for } r > 1 - d_{\text{a}}, \\ 1 & \text{else.} \end{cases}$$



Effects of plate separation (left) and influence of thickness in the astenosphere (right)



Surrogates in stochastic inversion



partly joint work with J.T. Oden, E. Lima, T. Yankeelov et.al. (2017,2018)
and with R. Scheichl and B. Gmeiner (2017)

Basics: Standard Multilevel Monte Carlo

- **Model problem:** $\nabla \cdot (k(x,\omega)\nabla p(x,\omega)) = f(x,\omega),\ \omega \in \Omega$

- **Sampling** from $k(x,\omega)$ by e.g.:
  - Truncated Karhunen-Loève (KL) expansion [Ghanem et al 91], [Widom 63]
    Circulant embedding, [Dietrich et al. 97, Graham et al. 18]
  - PDE-based variants, [Lindgren et al. 11]

- **Standard Monte Carlo** estimator: (Q is the quantity of interest)

$$\hat{Q}^{MC}_{N;h} := \frac{1}{N}\sum_{i=1}^{N} Q_h^{(i)}, \quad \text{MSE} = \frac{\mathbb{V}[Q_h]}{N} + (\mathbb{E}[Q_h - Q])^2$$

- **Standard Multilevel Monte Carlo** estimator: $h := h_L,\ h_l := 1/2h_{l-1},\ Q_{h_{-1}} := 0$

$$\text{MSE} = \sum_{l=0}^{L} \frac{\mathbb{V}[Q_{h_l} - Q_{h_{l-1}}]}{N_l} + (\mathbb{E}[Q_h - Q])^2, \quad \text{[Giles 08, Barth et al. 11, Cliffe et al. 11]}$$

Fractional PDE-based sampling

Matérn covariance [Matérn 60], [Abramowitz et al. 65] can be sampled solving a PDE:

$$\left(\kappa^2 - \Delta\right)^{\alpha/2} Y = \sigma \mathcal{W}, \quad \text{on } \mathbb{R}^d, \qquad \text{[Lindgren et al. 11]}$$

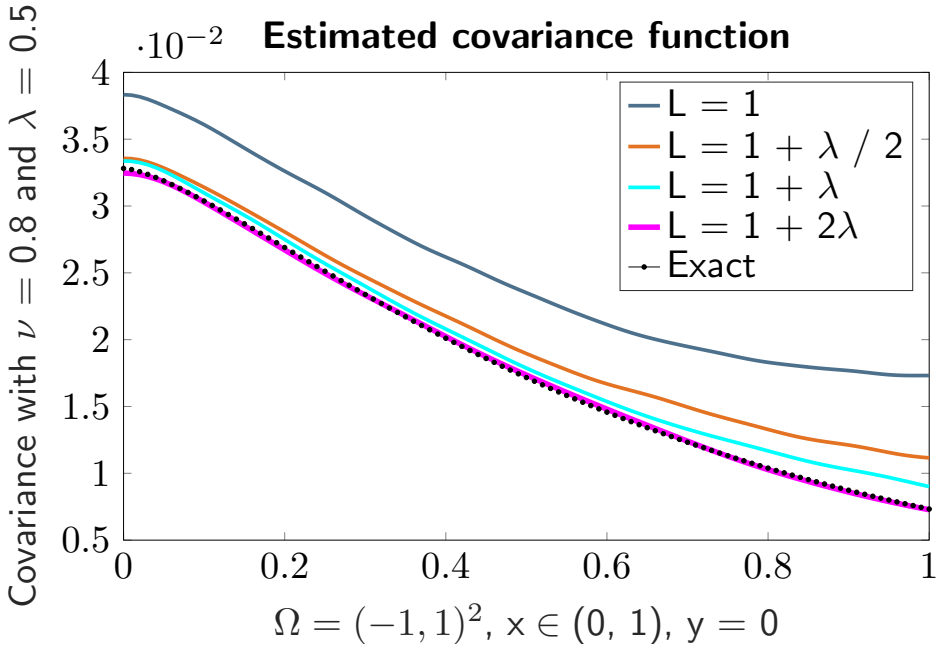
- $\kappa = 1/\lambda$ - inverse of the correlation length
- $\alpha = \nu + d/2$ - depends on the smoothness ν
- $\mathcal{W}$ - Gaussian white noise with mean zero and variance one

Window technique can be used to approximate Y on a bounded domain Ω

- Embed $\Omega \subset B_R^\infty(x_c) \subset B_L^\infty(x_c)$, with $L = R + k\lambda$
- Impose BCs on $B_L^\infty(x_c)$, e.g. hom. Neumann
- Approximate Y by Fourier techniques

Lemma: A priori estimate in terms of k:

$$\left|C_{Y_L}(x,y) - C_Y(x,y)\right| = \mathcal{O}\left(e^{-\beta k}\right)$$

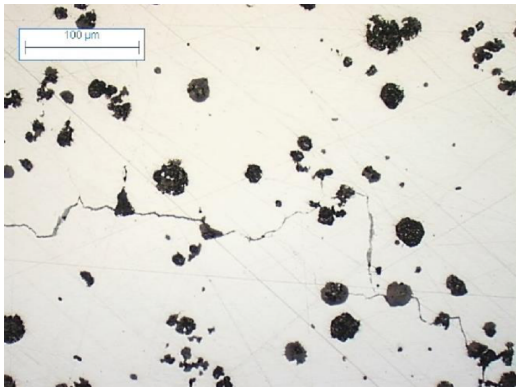


Sampling of synthetic two-phase material

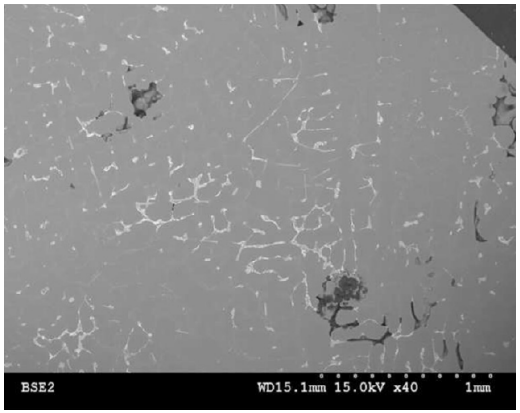
Two-phase material: χ_I (inclusions) and χ_M (matrix), ϕ volume fraction

$$\chi(\mathbf{x}) = \begin{cases} \chi_I, & \text{if } Y(\mathbf{x}) \geq \sqrt{2\,C(0)} \cdot \operatorname{erf}^{-1}(1 - 2\,\mathbb{E}\,[\phi]), \\ \chi_M, & \text{otherwise,} \end{cases}$$

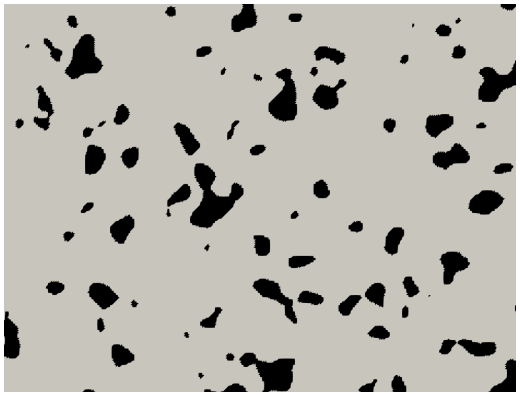
Cast iron with graphite inclusions [Szmytka et al., 17]



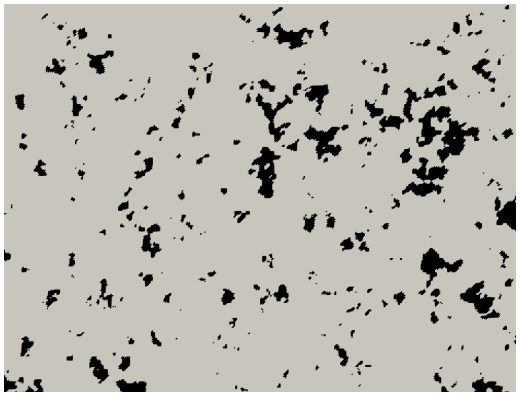
Al-Si alloy with pores inclusions [Charkaluk et al., 14]



$\nu = 10$



$\nu = 0.5$

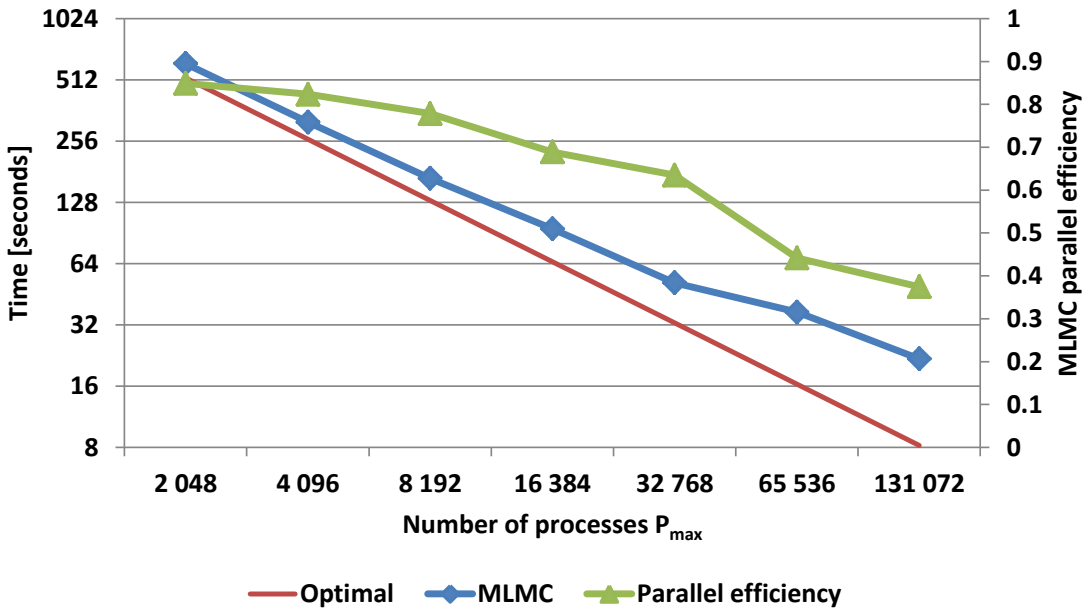


Scaling results for adaptive MLMC

- **Weak scaling** for adaptive MLMC

Cores	Mesh	Runtime	No. Samples Fine	No. Samples Total	Correlation length	Idle time
2 048	$1\,024^3$	$5.0 \cdot 10^3$ s	68	13 316	1.50E-02	3%
16 384	$2\,048^3$	$3.9 \cdot 10^3$ s	44	10 892	7.50E-03	4%
131 072	$4\,096^3$	$5.2 \cdot 10^3$ s	60	10 940	3.75E-03	5%

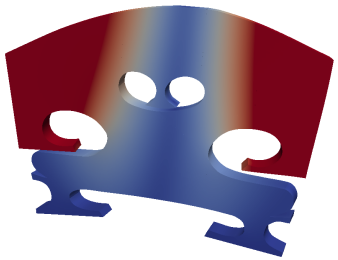
- **Strong scaling** for fixed sample numbers



Goal-oriented adaptive surrogate construction

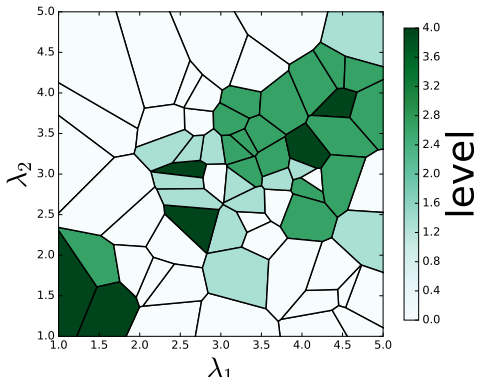
- **Different** types of refinement based on **approximation of the adjoint**:
  - **p-refinement:** local polynomial order of the surrogate on a Voronoi cell is increased
  - **level-refinement:** model level of the surrogate on a Voronoi cell is increased
  - **h-refinement:** new generating points for the Voronoi tessellation are added

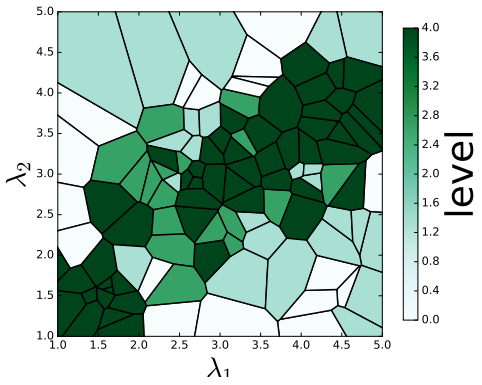
- **Nine** dimensional parameter space (orthotropic material parameters)

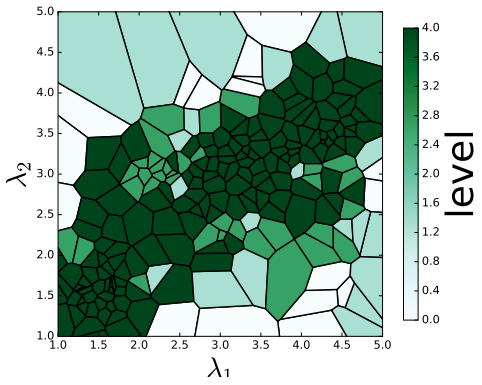


Level 1	Level 2	Level 3	Rel. Error	Run Time (s)
100	0	0	1.04e-02	6,544
102	70	0	5.69e-03	20,572
103	73	49	4.70e-04	35,708

- Simplified problem with **two** dimensional parameter space







Simplified avascular tumor model

Starting point: Mixture theory for different species α

$$\frac{\partial(\rho_\alpha \phi_\alpha)}{\partial t} + \text{div } (\rho_\alpha \phi_\alpha \mathbf{v}_\alpha) = \rho_\alpha(\text{div } \mathbf{J}_\alpha + S_\alpha)$$

$\mathbf{v}_\alpha$ convective velocity, ϕ_α volume fraction, ρ_α density, $\rho_\alpha \mathbf{J}_\alpha$ mass flux, S_α source term

Under additional assumptions, a simplified model can be obtained:

$$\begin{aligned}
\frac{\partial \phi_T}{\partial t} &= \text{div } (M_T \nabla \mu) + \lambda_{\text{prol}} \phi_\sigma \phi_T (1 - \frac{1}{K}\phi_T) - \lambda_{\text{apop}} \phi_T \\
\mu &= \Psi'(\phi_T) - {\epsilon_T}^2 \Delta \phi_T \\
\frac{\partial \phi_N}{\partial t} &= \lambda_{VN} H(\sigma_{VN} - \phi_\sigma)(\phi_T - \phi_N)
\end{aligned}$$

K carrying capacity, λ_{apop} apotosis rate, λ_{prol} rate of cellular mitosis, σ_{VN} transition point, λ_{VN} transition rate, ϵ_T interaction length, $\Psi(\phi_T) = E_T \phi_T^2 (1 - \phi_T)^2$ double well potential with energy scale E_T, M_T mobility matrix

Seven parameters have to be calibrated plus additional **hyperparameters** for the noise

Adaptive calibration for C3A liver cancer cells

Setup 1: Treatment of cell cultures with Mitomycin C to inhibit proliferation
Reduction of the PDE system to an ODE of expotential decline with rate λ_{apop}

Setup 2: Nutrient rich environment (concentration of fetal bovine serum (FBS) 10%)
Reduction of the PDE system to an ODE of logistic type λ_{apop}, K, λ_{prol}

Setup 3: Nutrient poor environment (concentration below 10% of FBS)
Reduction of the PDE system to a coupled ODE system λ_{apop}, K, λ_{prol}, λ_{apop}, σ_{VN}

Setup 4: Tracking of cells treated by green fluorescent protein in a short time interval
Reduction of the PDE system to a simple phase-field system M_T, ϵ_T

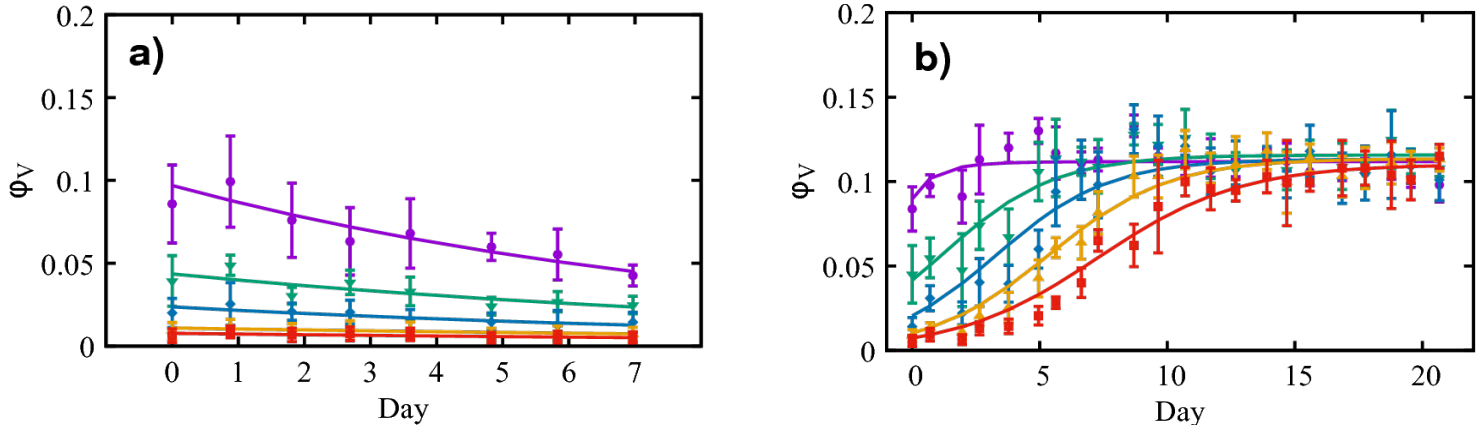
Bayesian update rule for the posterior: y experimental data, d model values

$$\pi(\theta|y) = \frac{\pi(y|\theta)\pi(\theta)}{\pi(y)}, \qquad \pi(y|\theta) = \prod_{j=1}^{J}\prod_{n=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(y_{ij} - d_j(\theta))^2}{2\sigma^2}\right)$$

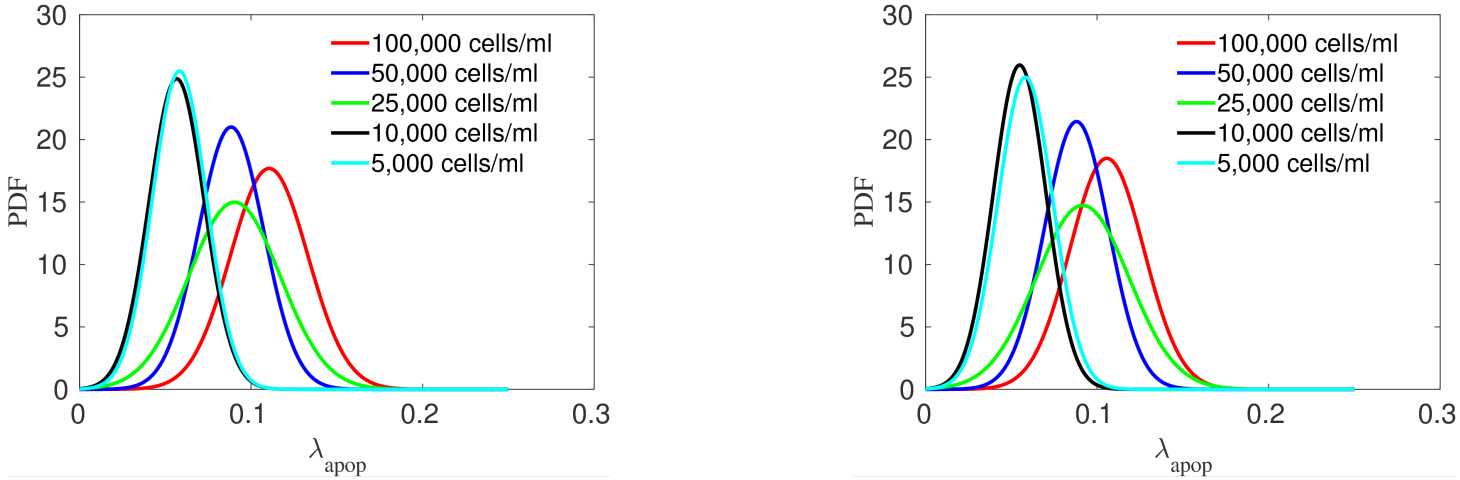
$\pi(\theta)$ prior probability density typical uniform or Gaussian (truncated)
$\pi(y)$ is a normalizing factor called model evidence, $\pi(y|\theta)$ likelihood function

Numerical results

- **Simulation** results versus experimental results for setup 1 (a) and setup 2 (b)



- **Posterior** for λ_{apop} after setup 1 (left) and setup 2 (right) (prior $\mathcal{U}(0, 10)$)



Experimental data is **informative** for the parameters in the simplified sub-models

Summary and conclusion

- Modern architectures **require** reevaluation of performance
- **Need** for surrogate operators in large scale FE simulations
- **Need** for surrogate models in complex applications
- Calibration in case of uncertainties **benefits** from hierarchical strategies