Chair of Scientific Computing
School of Computation, Information and Technology
Technical University of Munich

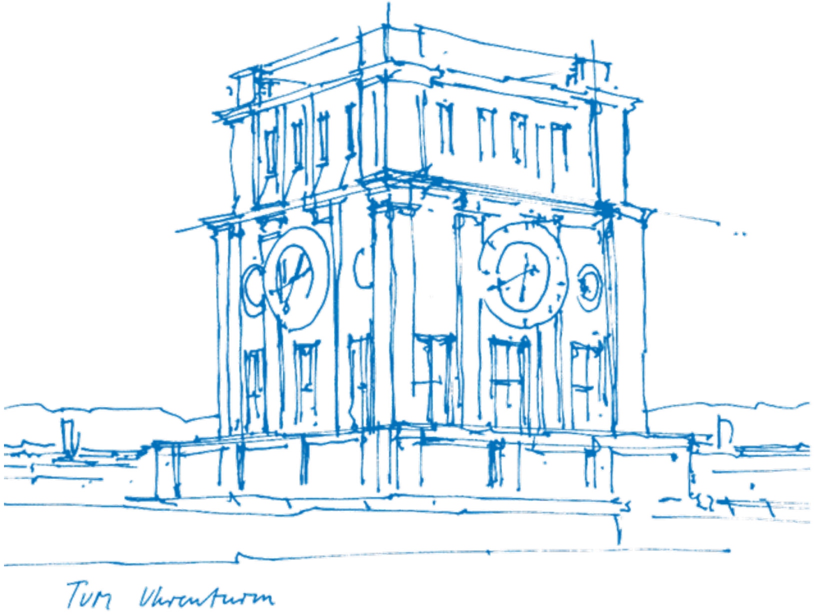
Dimension adaptive Multi-fidelity Polynomial Chaos Expansion using Leja points

Sparse Grids and Applications Seminar 2024

Kislaya Ravi , Tobias Neckel, Hans-Joachim Bungartz

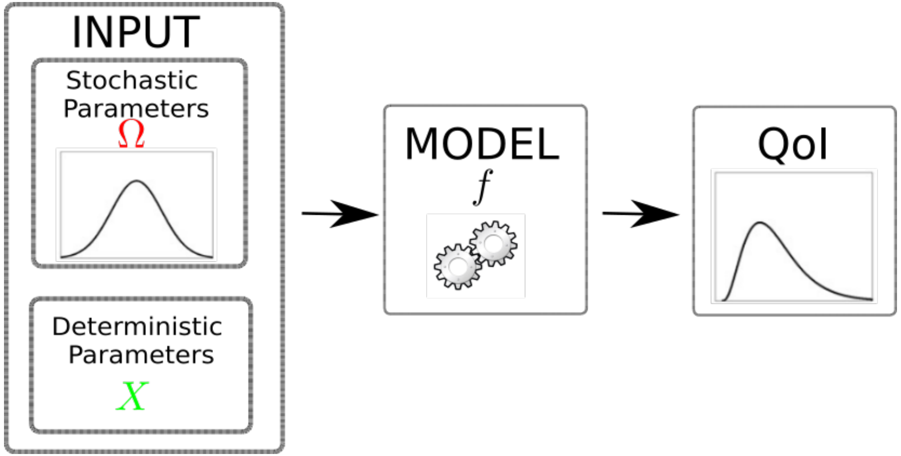
Technical University of Munich
School of Computation, Information and Technology
Chair of Scientific Computing

Bonn, 11. September, 2024

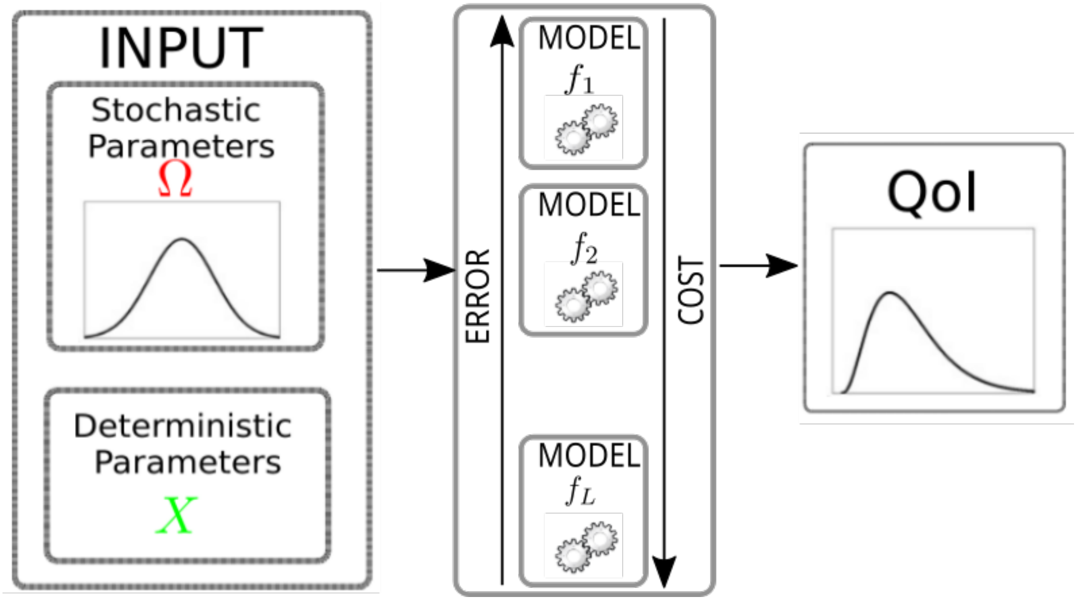


Forward UQ: Problem Statement

- **Given:** A function $f(X, \Omega)$, where
  - X are deterministic parameters
  - Ω are stochastic parameters
- For the given $t \in X$ and distribution of $\omega \in \Omega$, what will be the distribution of our Quantity of interest (QoI)



Multi-fidelity Forward UQ: Problem Statement



Polynomial Chaos Expansion

- approximate $f(t,\omega)$ by series of polynomials

$$f(t,\omega) \approx \mathscr{F}(t,\omega) = \sum_{n=0}^{\infty} \hat{f}_n(t)\, \phi_n(\omega)$$

- $\phi_n(\omega)$ orthonormal polynomials of degree n, $\hat{f}_n(t)$ coefficients

- Truncate the series to N terms

$$\mathscr{F}(t,\omega) = \sum_{n=0}^{N-1} \hat{f}_n(t)\, \phi_n(\omega)$$

- In higher dimension, we take the tensor product of the polynomials.
- Pseudo-spectral approach to calculate $\hat{f}_n$

$$\hat{f}_n(t) = \int_{\Omega} f(t,\omega)\, \phi_n(\omega)\, \rho(\omega)\, d\omega = \sum_{k=1}^{K} w_k \phi(x_k) f(t, x_k)$$

- Statistical moments can be easily calculated as:

$$\mathbb{E}[f(t,\omega)] = \hat{f}_0$$

$$\mathbb{V}[f(t,\omega)] = \sum_{i=1}^{N-1} \hat{f}_i^2$$

Challenges

- Determine the order of Polynomial $\implies$ Dimension Adaptivity[1]
- Curse of Dimensionality $\implies$ Sparse grid
- Minimize number of function evaluations $\implies$ Leja Points
- Multi-fidelity $\implies$ Correction terms[2]

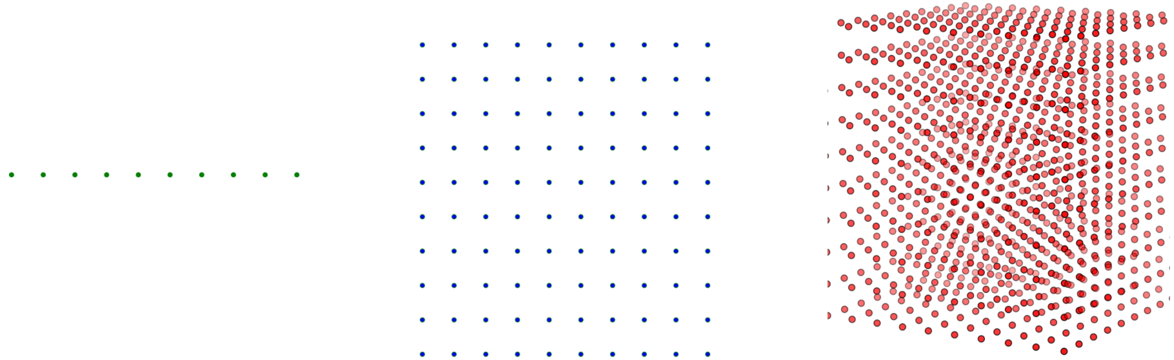
---

[1] Ionuţ-Gabriel Farcaş et al. "Sensitivity-driven adaptive sparse stochastic approximations in plasma microinstability analysis". In: *Journal of Computational Physics* (2020), p. 109394.

[2] Leo Wai-Tsun Ng and Michael Eldred. "Multifidelity uncertainty quantification using non-intrusive polynomial chaos and stochastic collocation". In: *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*. 2012, p. 1852.

Sparse Grid

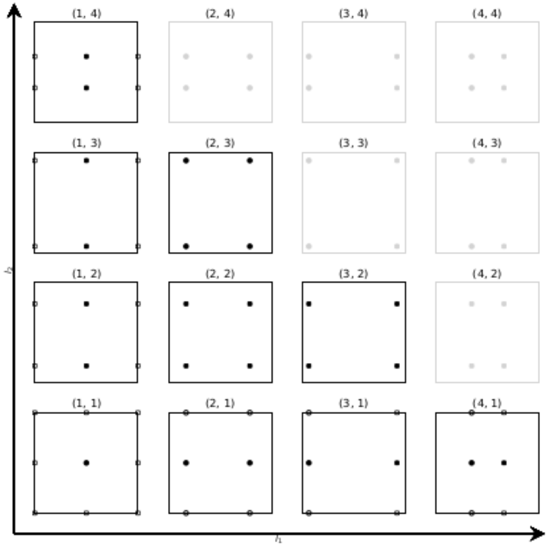
- Full tensor-grid approach assumes that all the directions are equally well coupled

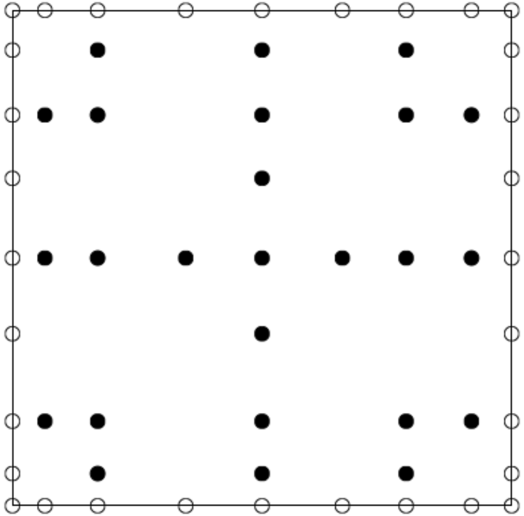


- **Idea:** Weaken the assumed coupling
- Discard the components that have low contribution to the overall solution

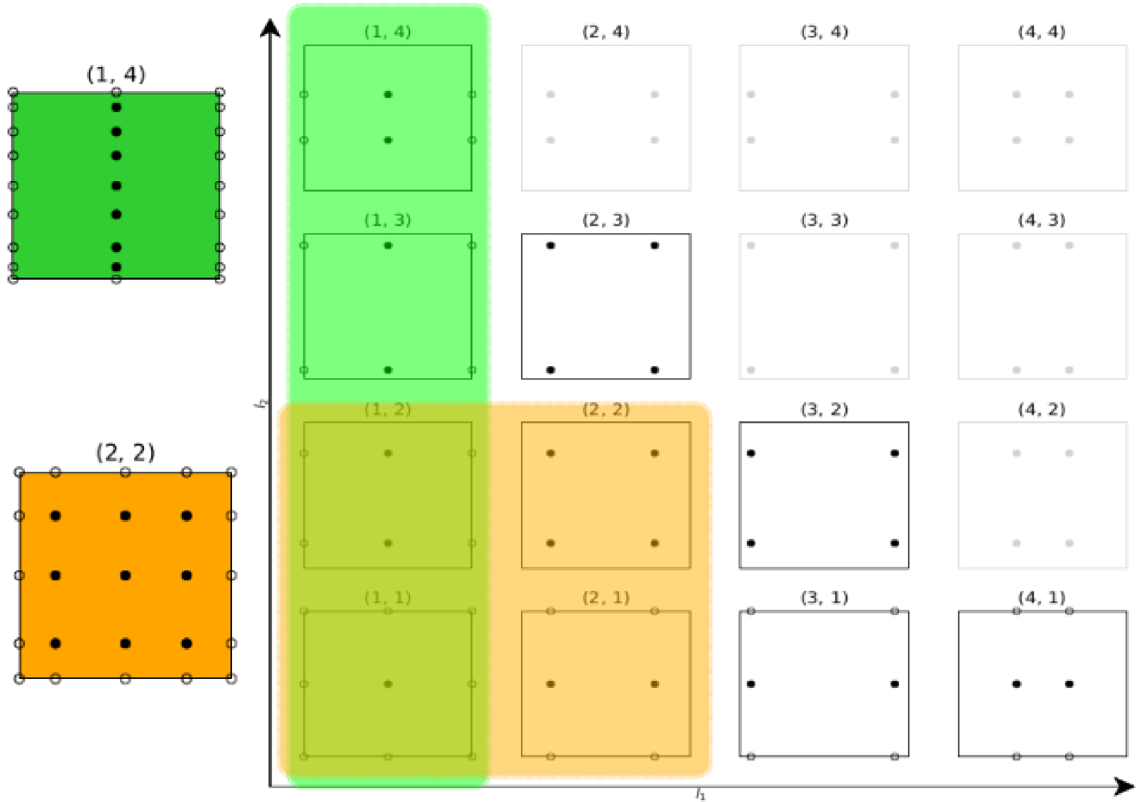
Multiindex

- Additional grid points at each level.
- We represent each set by a *Multiindex*
- Ignore higher order terms: But which ones?





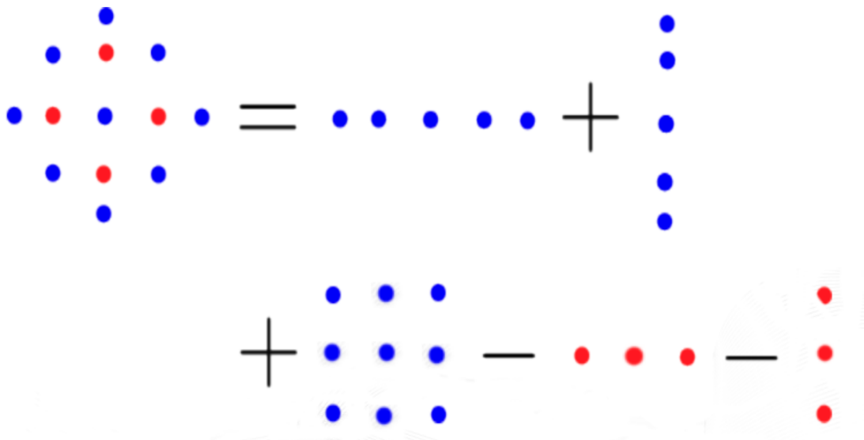
Combination of multiindex



Combine Grids

Any grid pattern can be written as linear combination of other grid patterns.[3,4]

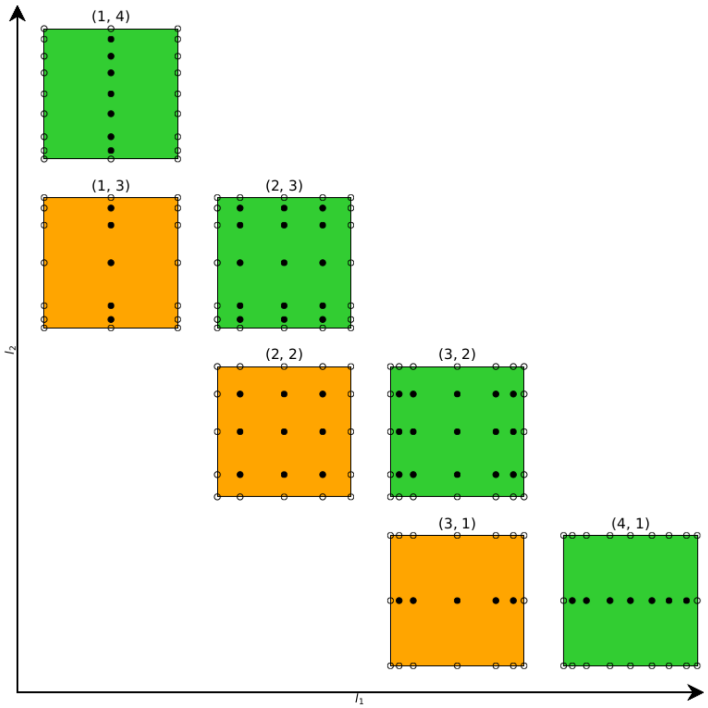
For example :

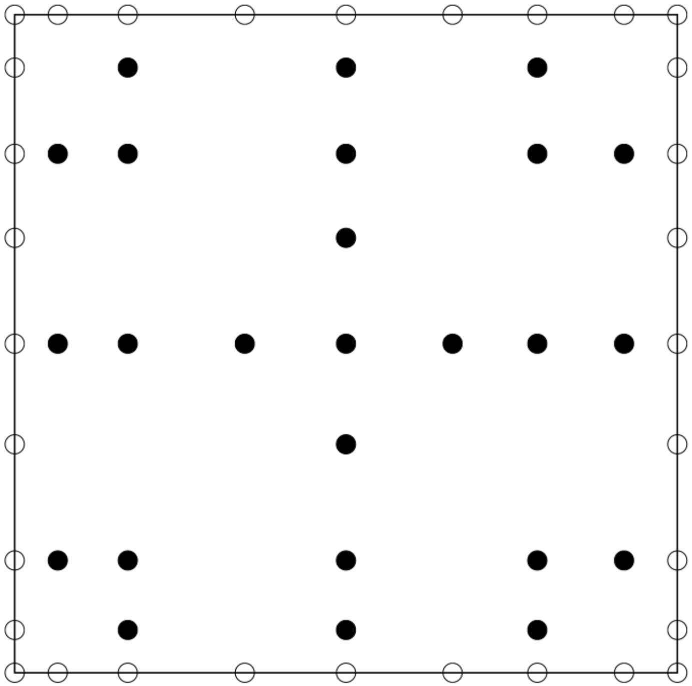


[3] Sergei Abramovich Smolyak. “Quadrature and interpolation formulas for tensor products of certain classes of functions”. In: *Doklady Akademii Nauk*. Vol. 148. 5. Russian Academy of Sciences. 1963, pp. 1042–1045.

[4] Michael Griebel, Michael Schneider, and Christoph Zenger. “A combination technique for the solution of sparse grid problems”. In: (1990).

Combination Technique





Dimension Adaptivity

- Adaptive choice of multiindex[a]

Algorithm 1: Single fidelity Adaptive Sparse Grid Approximation: version 1

input: Stochastic dimension (d), number of adaption steps (N_S), function (f)

output: Set of multiindices $\mathcal{A}$

$\mathcal{A} := \{\mathbb{1}_d\}$;

for $n \leftarrow 1$ ***to*** N_S **do**

  $\mathcal{O} := \{a | a - e_i \in \mathcal{A}, \forall i = 1, 2...d\}$;

  foreach $o \in \mathcal{O}$ **do**

    Evaluate PCE($\mathscr{F}_o$) coefficients $\hat{f}^o$;

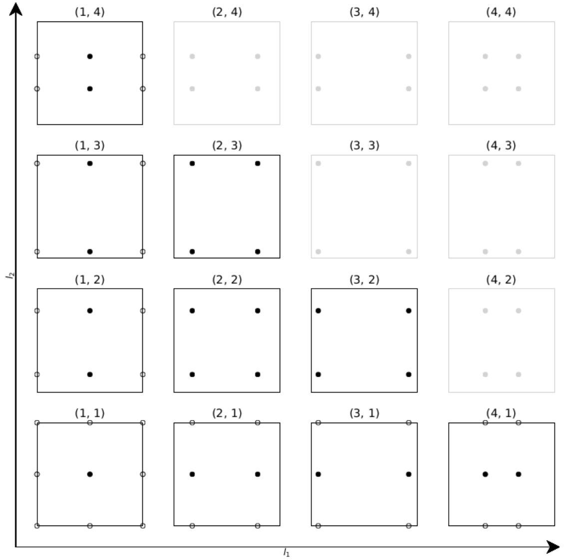
    $\Delta^o := \mathbb{V}[\mathscr{F}_{\mathcal{A}\cup o}] - \mathbb{V}[\mathscr{F}_{\mathcal{A}}] + (\mathbb{E}[\mathscr{F}_{\mathcal{A}\cup o}] - \mathbb{E}[\mathscr{F}_{\mathcal{A}}])^2$;

  end

  $s := \underset{o \in \mathcal{O}}{\operatorname{argmax}} \Delta^o$;

  $\mathcal{A} := \mathcal{A} \cup s$;

end



[a] Farcaş et al., "Sensitivity-driven adaptive sparse stochastic approximations in plasma microinstability analysis", op. cit.

Surplus Calculation

- Surplus is the sum of change in variance and square of change in mean the due to addition of multiindex o.
- This measures the contribution of the multiindex o to the overall solution.
- Surplus depends upon the coefficients of the PCE.
- Addition of multiindex only effects the neighbors
- So, the change in coefficients due to addition of multiindex o for combination of polynomial order n ($\Delta\hat{f}_n^o$)

$$\Delta\hat{f}_n^o = \sum_{z\in\{0,1\}^d} (-1)^{|z|_1}\hat{f}_n^{o-z}$$

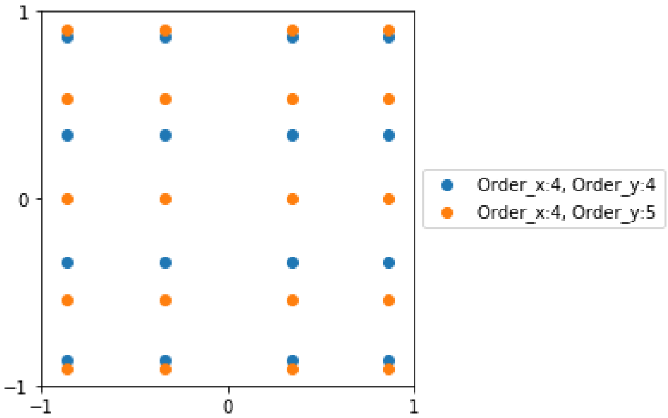
- Variance surplus is

$$\Delta^o = \sum_{i\in\mathcal{A}\cup o} (\Delta\hat{f}_i^o)^2 - 2\Delta\hat{f}_i^o\hat{f}_i^{\mathcal{A}}$$

Choice of quadrature points

Issues

- During adaptivity we realize that a higher order polynomial is required
- This will need all new sets of quadrature points
- Number of quadrature points grows exponentially for many methods like gaussian quadrature etc.



So, we need points that are:

- Nested
- Spawn less points per level

Leja Points[5]

$$\theta_{2l} = \underset{\theta \in \Omega}{\operatorname{argmax}} \left| \prod_{l=1}^{2n-1} (\theta - \theta_l) \right| \rho(\theta)$$

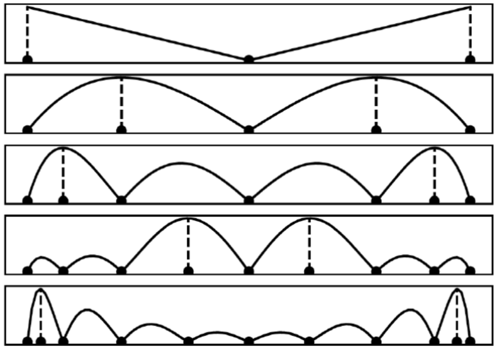


Figure: 1D Leja points per level

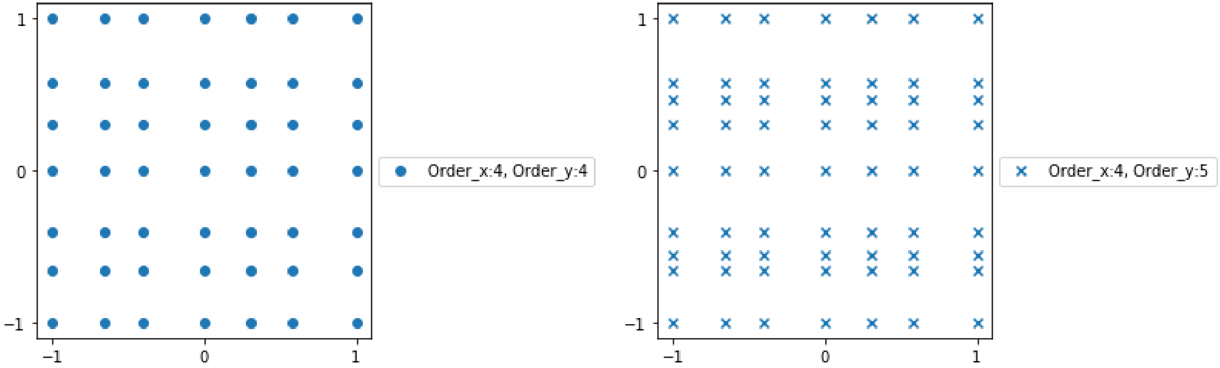
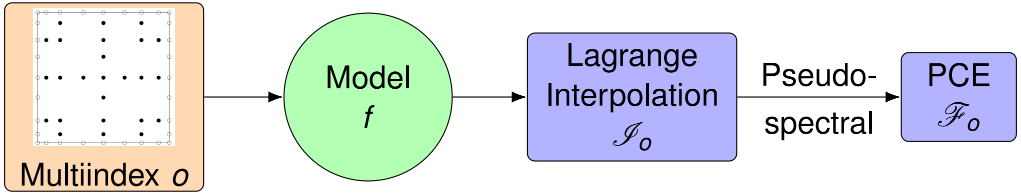


Figure: Growth of points in 2D

[5] Peter Jantsch, Clayton G Webster, and Guannan Zhang. "On the Lebesgue constant of weighted Leja points for Lagrange interpolation on unbounded domains". In: *IMA Journal of Numerical Analysis* 39.2 (2019), pp. 1039–1057.

Middle layer of Interpolation

- Leja points are suited for interpolation.[6]
- Instead of using Leja points for quadrature, we use them for interpolation.
- Create PCE over the interpolated function.



[6] Akil Narayan and John D Jakeman. "Adaptive Leja sparse grid constructions for stochastic collocation and high-dimensional approximation". In: *SIAM Journal on Scientific Computing* 36.6 (2014), A2952–A2983.

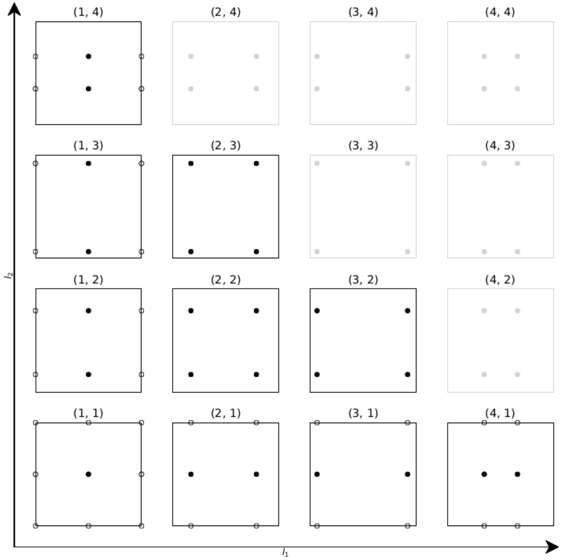
Algorithm revisited

Algorithm 2: Single fidelity Adaptive Sparse Grid Approximation[a]

input: Stochastic dimension (d), number of adaption steps (N_S), function (f)
output: Set of multiindices $\mathcal{A}$

$\mathcal{A} := \{\mathbb{1}_d\}$;
for $n \leftarrow 1$ ***to*** N_S **do**
  $\mathcal{O} := \{a | a - e_i \in \mathcal{A}, \forall i = 1, 2 ... d\}$;
  foreach $o \in \mathcal{O}$ **do**
    Build $\mathscr{I}_o$ using Leja points ;
    Evaluate PCE($\mathscr{F}_o$) coefficients $\hat{f}^o$ using $\mathscr{I}_o$;
    $\Delta^o := \mathbb{V}[\mathscr{F}_{\mathcal{A} \cup o}] - \mathbb{V}[\mathscr{F}_{\mathcal{A}}] + (\mathbb{E}[\mathscr{F}_{\mathcal{A} \cup o}] - \mathbb{E}[\mathscr{F}_{\mathcal{A}}])^2$;
  end
  $s := \underset{o \in \mathcal{O}}{\operatorname{argmax}} \Delta^o$;
  $\mathcal{A} := \mathcal{A} \cup s$;
end

[a] Farcaş et al., "Sensitivity-driven adaptive sparse stochastic approximations in plasma microinstability analysis", op. cit.



Multi-fidelity

- Express the high fidelity function (f_h) as sum of low fidelity function (f_l) and a correction term(δ)[7]

$$f_h = \gamma(f_l + \delta_d) + (1 - \gamma) f_l \delta_r$$

where

$$\delta_d = f_h - f_l \qquad \delta_r = \frac{f_h}{f_l}$$

- γ is obtained by minimising the surplus term

$$\gamma = \frac{\Delta^2_{\delta_r}}{\Delta^2_{\delta_d} + \Delta^2_{\delta_r}}$$

- Apply Algorithm 1 with QoI as:
  - f_h to get low fidelity multiindex
  - δ to get high fidelity multiindex
- Out of the two fidelities select the one with higher contribution to overall variance.

[7] Ng and Eldred, "Multifidelity uncertainty quantification using non-intrusive polynomial chaos and stochastic collocation", op. cit.

Multi-fidelity Algorithm

Algorithm 3: Multi-fidelity dimension adaptive sparse grid

input: Stochastic dimension (d), number of adaption steps (N), function ($f_l, \quad l = 1, 2, ...L$)

output: Set of multi-indices $\mathcal{A}_l \quad l = 1, 2, ..., L$

$\mathcal{A}_l := \{\mathbb{1}_d\}, l = 1, 2, .., L$;

for $n \leftarrow 1$ ***to*** N **do**

  for $l \leftarrow 1$ ***to*** L **do**

    $\mathcal{O}_l := \{a | a - e_i \in \mathcal{A}_l, \forall i = 1, 2...d\}$;

    foreach $o \in \mathcal{O}_l$ **do**

      $\Delta_o^l := \mathbb{V}[f]_{\mathcal{A}_l \cup o} - \mathbb{V}[f]_{\mathcal{A}_l} + (\mathbb{E}[f]_{\mathcal{A}_l \cup o} - \mathbb{E}[f]_{\mathcal{A}_l})^2$;

    end

  end

  $q, s := \underset{l=1,..,L; o \in \mathcal{O}_l}{\operatorname{argmax}} \Delta_o^l$;

  $\mathcal{A}_q := \mathcal{A}_q \cup s$;

end

Results

- Academic benchmark
- Multi-fidelity with different mesh size (multi-level)
- Multi-fidelity with different undelying modeling assumptions

Academic benchmark

Product of sinusoidal function

$$f_h(X) = \prod_{i=1}^{d} \sin a_i x_i$$

$$f_l(X) = f_h(X) + g(X)$$

where, $X \in \mathbb{R}^d, a_i, \in \mathbb{R}, x_i \sim \mathscr{U}[0,1], i = 1,2,....,d, d \in \mathbb{N}, g : \mathbb{R}^d \to \mathbb{R}$. The analytical mean and variance of f is

$$\mathbb{E}[f_h] = \prod_{i=1}^{d} \frac{1-\cos a_i}{a_i}$$

$$\mathbb{V}[f_h] = \left(\prod_{i=1}^{d}\left(0.5 - \frac{\sin 2a_i}{4a_i}\right)\right) + (\mathbb{E}[f_h])^2 + \left((-1)^d \times 2\mathbb{E}[f_h]\prod_{i=1}^{d}\left(\frac{\cos a_i - 1}{a_i}\right)\right)$$

Academic bencmark

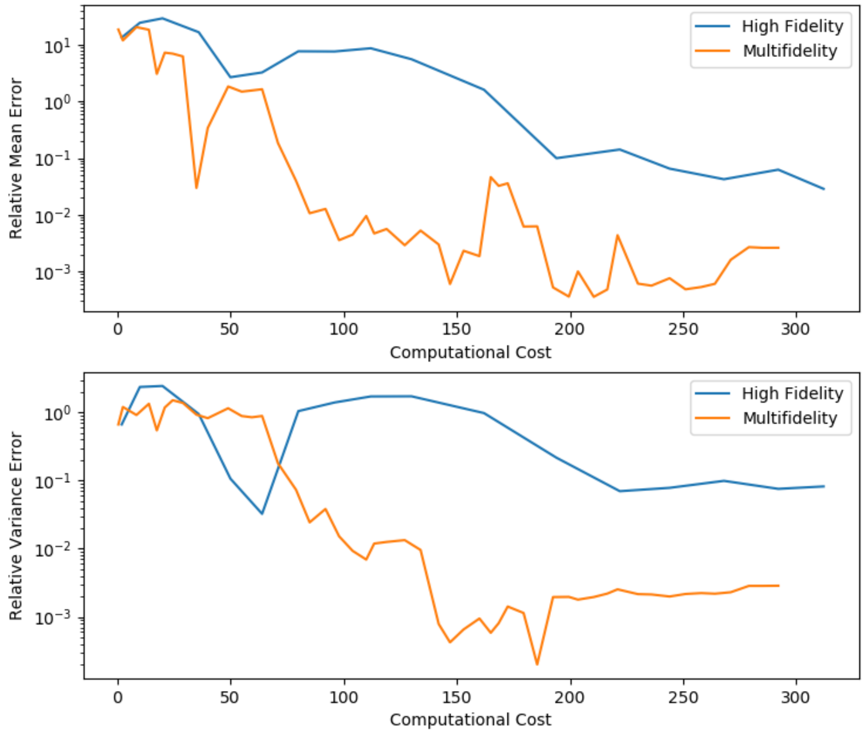
Problem Statement

$$f_h(X) = \sin(\pi x_1) \sin\left(\frac{3\pi x_2}{2}\right) \sin\left(\frac{5\pi x_3}{2}\right)$$

$$f_l(X) = f_h(X) + \sin\left(\frac{x_1}{2}\right) + \sin\left(\frac{3x_2}{4}\right) + \sin\left(\frac{x_3}{2}\right)$$

We assume that high fidelity function takes 4 times more time than the low fidelity function

Academic benchmark



Poisson Equation

Problem Statement

- Elliptic PDE (Six dimensional example)
- Consider a stochastic PDE in two dimensional spatial domain

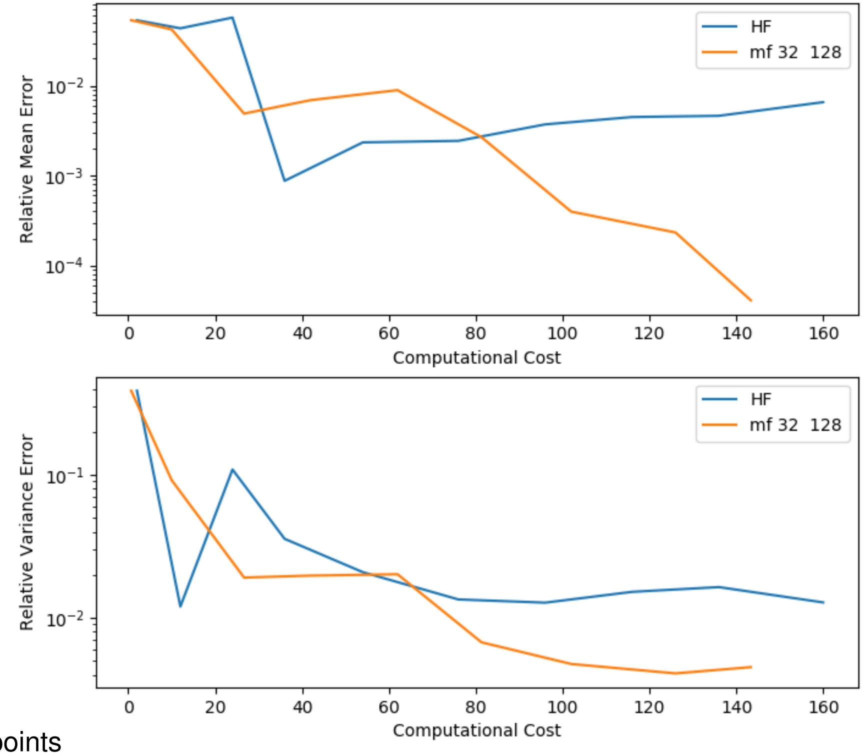
$$-\frac{\partial}{\partial x}\left[\kappa(x,\omega)\frac{\partial u(x,\omega)}{\partial x}\right] = g(x), \quad x \in [0,1]^2$$

- Zero Dirichlet boundary condition
- The diffusion coefficient is described by 6-dimensional Karhunen-Loeve expansion

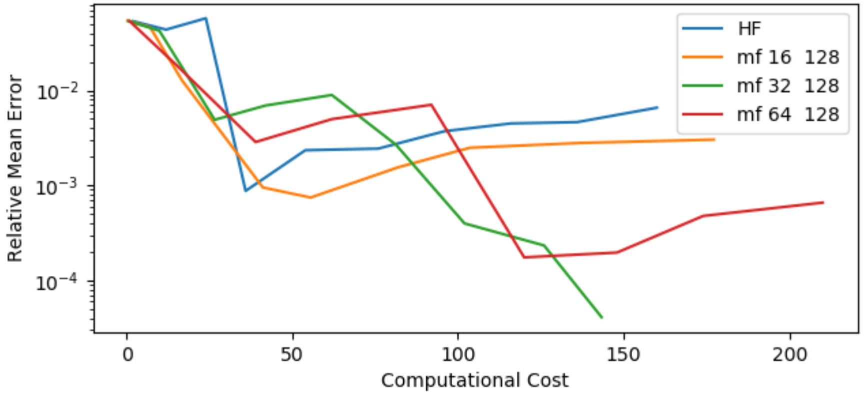
$$\kappa(x,\omega) = 1.5 + \sqrt{2}\sum_{i=1}^{6}\sqrt{\lambda_k}\Phi_k(x)Y_k(x), \quad Y_k \sim \mathscr{U}[0,1]$$

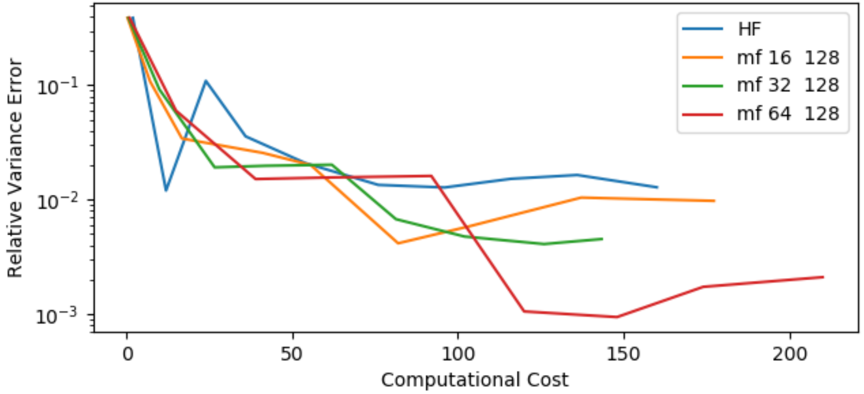
- Φ and λ are eigen vector and value for exponential co-variance kernel with correlation length 1.
- Fidelity depends upon the mesh size of the FEM solver.

Poisson Equation: Results



Poisson Equation: Results





Transport Analysis of Tokamak Experiments

- We use **A**utomated **S**ystem for **TR**ansport Analysis (ASTRA)[a]
- We use Quasilinear transport model with saturation rules derived from Gyrokinetic codes.
- We use Qualikiz[b] as high fidelity model.
- We use QLKNN[c] as low fidelity. This is physics based neural network trained on 3×10^7 data points. It is 10^4 times faster than Qualikiz.

[a] Gregorij V Pereverzev and PN Yushmanov. "ASTRA. Automated System for TRansport Analysis in a tokamak". In: (2002).

[b] C Bourdelle et al. "Core turbulent transport in tokamak plasmas: bridging theory and experiment with QuaLiKiz". In: *Plasma Physics and Controlled Fusion* 58.1 (2015), p. 014036.

[c] Karel Lucas van de Plassche et al. "Fast modeling of turbulent transport in fusion plasmas using neural networks". In: *Physics of Plasmas* 27.2 (2020), p. 022310.

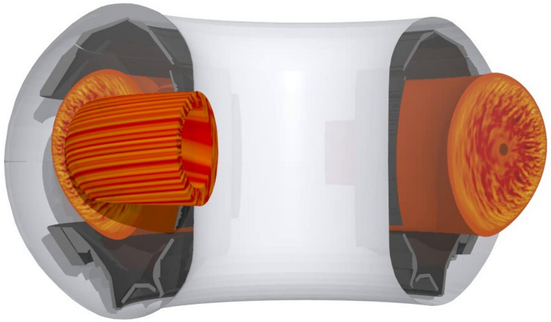


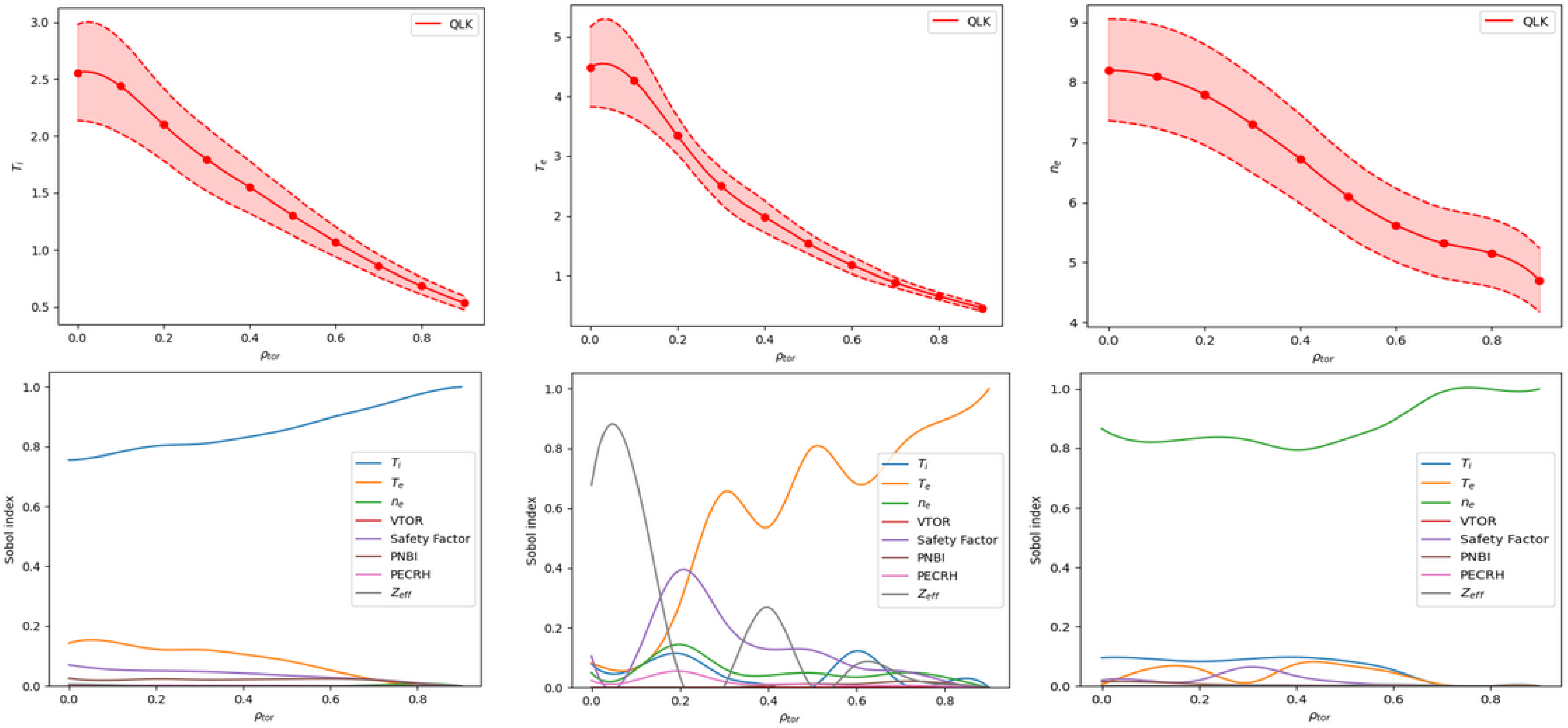
Figure: Section of Tokamak reaction[a]

[a] Tobias Goerler et al. "The global version of the gyrokinetic turbulence code GENE". In: *Journal of Computational Physics* 230.18 (2011), pp. 7053–7071.

Transport Analysis of Tokamak Experiments

- List of uncertain parameters are:
  - Initial ion temperature measurements (T_i)
  - Initial electron temperature measurements (T_e)
  - Initial electron density measurements (n_e)
  - Toroidal rotation ($VTOR$)
  - Safety factor
  - Effective charge (Z_{eff})
  - NBI heating
  - ECRH heating
- We assume uniform distribution within $\pm 10\%$ of the experimental value.
- We choose shot number 33616, with 5MW of NBI heating and 1.2 MW of ECRH heating.

Transport Analysis of Tokamak Experiments



Conclusion and Future works

Conclusion:

- Saved computational resources by employing Multi-fidelity framework along with Leja points.
- Results depend on the quality of the low fidelity method.
- Complex non-linear relationship is difficult to model.

Future Works:

- Model high fidelity function as composite function:

$$f_h(X) = g(f_l(X), X)$$

References

- Bourdelle, C et al. “Core turbulent transport in tokamak plasmas: bridging theory and experiment with QuaLiKiz”. In: *Plasma Physics and Controlled Fusion* 58.1 (2015), p. 014036.
- Farcaş, Ionuţ-Gabriel et al. “Sensitivity-driven adaptive sparse stochastic approximations in plasma microinstability analysis”. In: *Journal of Computational Physics* (2020), p. 109394.
- Goerler, Tobias et al. “The global version of the gyrokinetic turbulence code GENE”. In: *Journal of Computational Physics* 230.18 (2011), pp. 7053–7071.
- Griebel, Michael, Michael Schneider, and Christoph Zenger. “A combination technique for the solution of sparse grid problems”. In: (1990).
- Jantsch, Peter, Clayton G Webster, and Guannan Zhang. “On the Lebesgue constant of weighted Leja points for Lagrange interpolation on unbounded domains”. In: *IMA Journal of Numerical Analysis* 39.2 (2019), pp. 1039–1057.
- Narayan, Akil and John D Jakeman. “Adaptive Leja sparse grid constructions for stochastic collocation and high-dimensional approximation”. In: *SIAM Journal on Scientific Computing* 36.6 (2014), A2952–A2983.
- Ng, Leo Wai-Tsun and Michael Eldred. “Multifidelity uncertainty quantification using non-intrusive polynomial chaos and stochastic collocation”. In: *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*. 2012, p. 1852.
- Pereverzev, Gregorij V and PN Yushmanov. “ASTRA. Automated System for TRansport Analysis in a tokamak”. In: (2002).
- Plassche, Karel Lucas van de et al. “Fast modeling of turbulent transport in fusion plasmas using neural networks”. In: *Physics of Plasmas* 27.2 (2020), p. 022310.
- Smolyak, Sergei Abramovich. “Quadrature and interpolation formulas for tensor products of certain classes of functions”. In: *Doklady Akademii Nauk*. Vol. 148. 5. Russian Academy of Sciences. 1963, pp. 1042–1045.

Thank You

Questions and Feedbacks