What is the added value of traditional methods for physics modelling?

Kathrin Smetana

(University of Twente)

Collaborators: A. T. Patera (M.I.T.), O. Zahm (INRIA), C. Brune (UT)

Outline

- Brief outline how we obtain predictions based on physics-based equations to illustrate …
- added value of physics-based modelling such as
  - stability, robustness, well-posedness
  - assessment of accuracy via error estimators
- How to combine traditional methods and machine learning

Making predictions based on physics-based equations

- Step 1: Modeling: Describe phenomena with physics-based equations (ordinary or partial differential equations (PDE)) on a certain domain.

- Step 2: Approximation: Use for instance Finite Element Method to discretize PDE. Results in linear system of equations we have to solve.

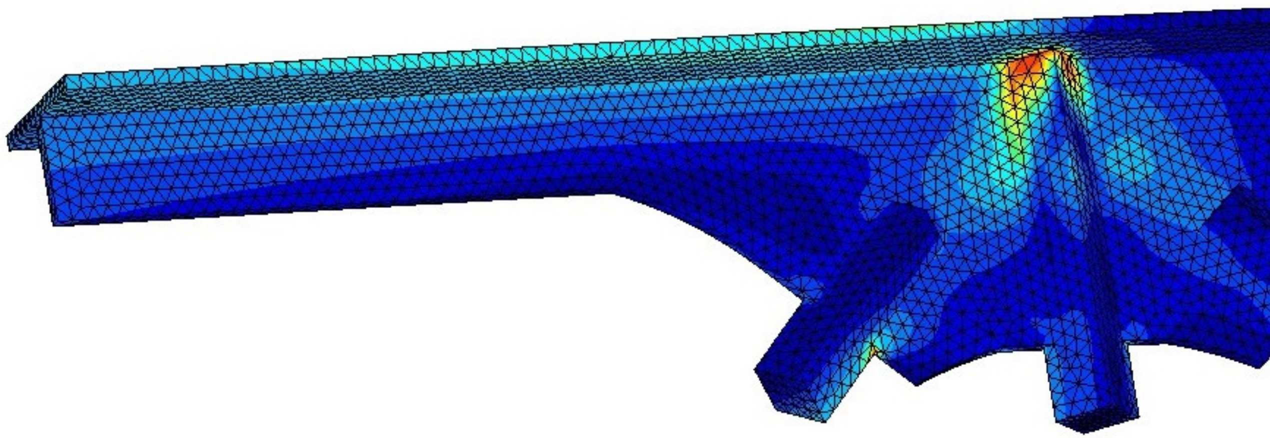
- Step 3: Acceleration: Fast solvers, reduced order modelling,…

Example:

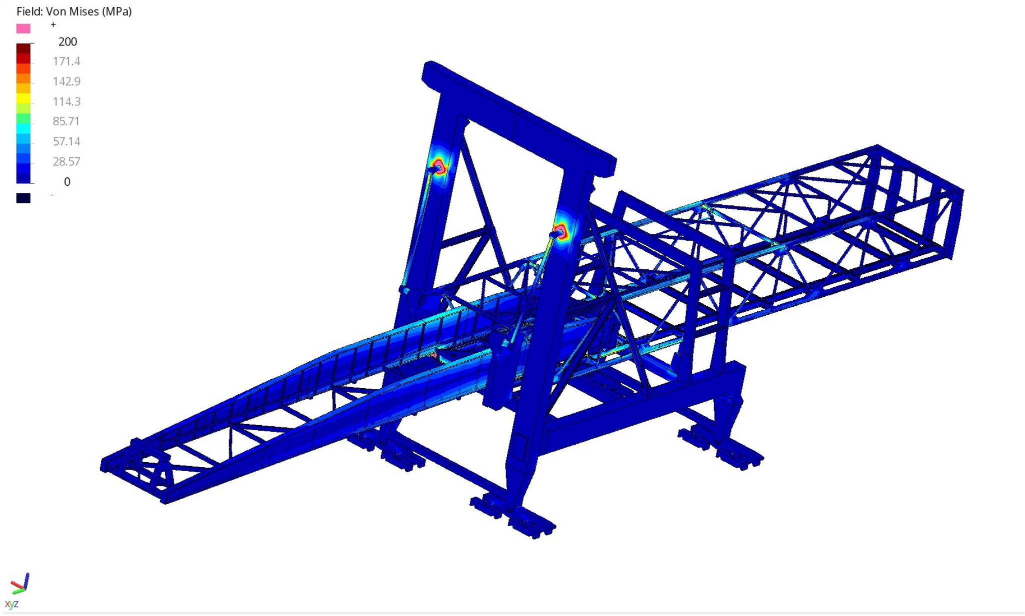
- Equations of linear elasticity: Find the displacement vector u and the Cauchy stress tensor $\sigma(\mathrm{u})$ such that

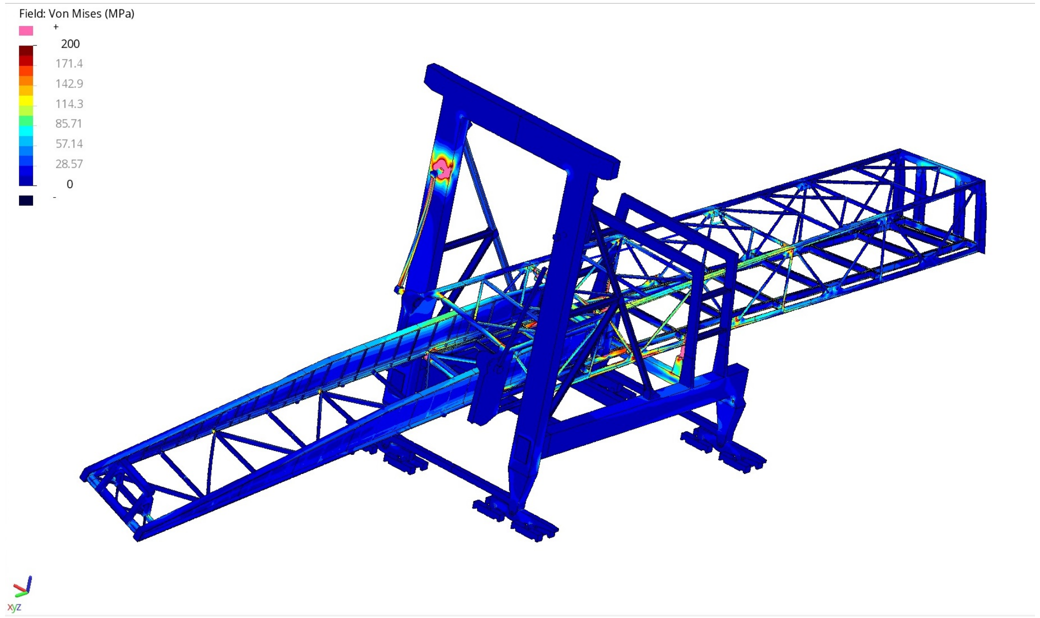
$$-\nabla \cdot \sigma(u) = f$$ + boundary conditions

- Find U that satisfies $AU = F$.



Making predictions based on physics-based equations





- FEM discretization: more than 20 millions degrees of freedom
- Dimension Schur complement: about 349 000

- Simulation time with reduced interface spaces: 2 seconds
- Dimension reduced Schur complement: about 12 000

Results on shiploader by company Akselos using reduced interface spaces introduced in K. Smetana, A.T. Patera, SIAM J. Sci. Comput., 2016.

Various sources of errors

- Model error (equations of linear elasticity do not describe phenomenon perfectly)
- Data error (measurements of data such as Young's modulus is prone to errors)
- Discretization error (error due to FEM approximation)
- Error due to acceleration (reduced model,…)
- Truncation error (error caused by linear systems of equations solver)

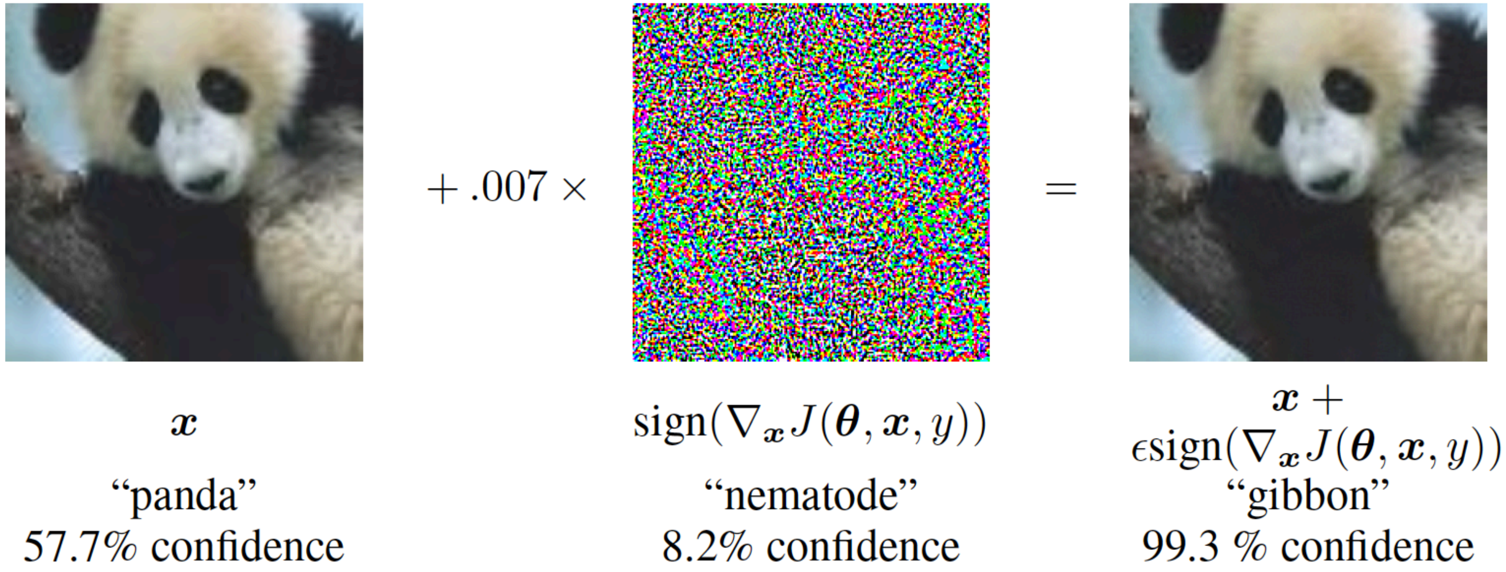
We have errors in every step, some are unavoidable

GOAL: Nevertheless ensure that we can relate prediction to true phenomenon.

Added value of physics-based modelling

1. Stability, robustness, well-posedness
2. Accuracy can be assessed and analyzed, for instance, by a posteriori or a priori error bounds
3. We are in general able to interpret, understand, and explain the results.

Stabilization issues with Deep Nets



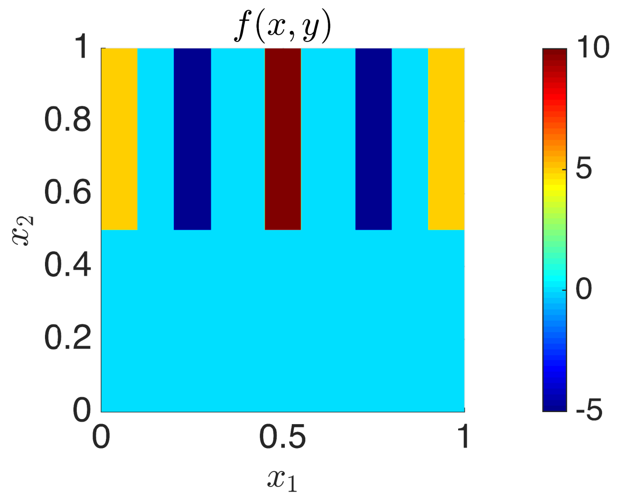
- Small changes in input data can have a significant effect
- Related problem: observation of vanishing or exploding gradients

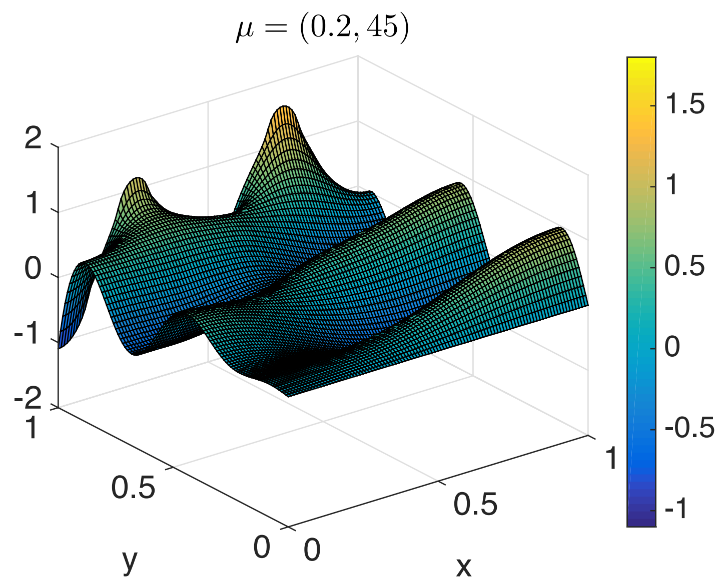
I. Goodfellow, J. Shlens, C. Szegedy, CoRR 2015, A. Nguyen, J. Yosiniski, J. Clune, In Computer Vision and Pattern Recognition (CVPR '15), IEEE, 2015, Antun et al., arXiv: 1902:05300, Y. Bengio, P. Simard, and P. Frasconi, IEEE Transactions on Neural Networks, 1994.

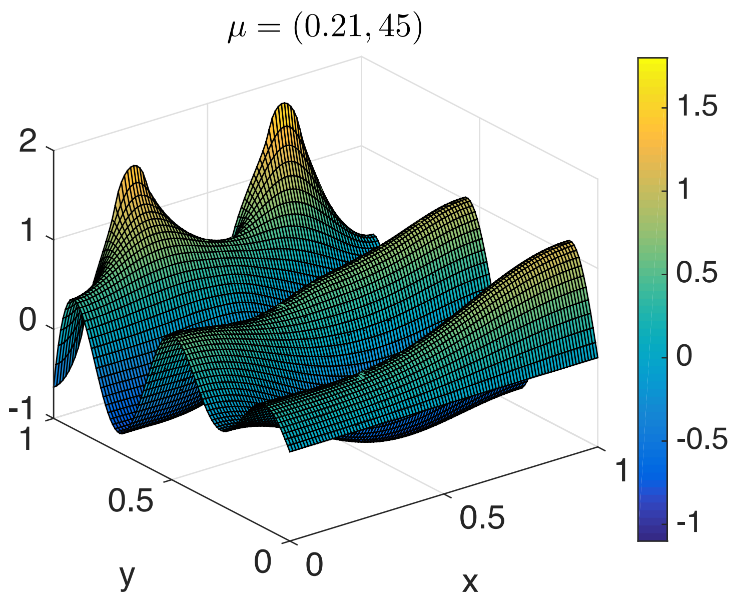
Stability in the context of physics-based modelling

- Consider anisotropic Helmholtz equation:

$$-\partial_{x_1x_1}u(x;\mu)-\mu_1\partial_{x_2x_2}u(x;\mu)-\mu_2u(x;\mu)=f(x) \quad \text{in } D \; + \text{b.c.}$$







- We have: $\|u(\mu)-u(\tilde{\mu})\|_{H^1(D)} \leq C\|\mu-\tilde{\mu}\|_2$ (Stability!)

Stability in the context of physics-based modelling

- Consider for instance $-div(a\nabla u) = f \;\; in \;\; D$. Then we have

$$\|u_{\tilde{a}} - u_a\|_{H_0^1(D)} \leq C\|a - \tilde{a}\|_{L^\infty(D)}$$

For instance: A. Bonito et al, SIAM J. Math. Anal., 2017.

- Similarly for the nonlinear PDE $\quad A(u) = f \quad in\ D \quad$ we obtain under certain verifiable conditions

$$\|u_{\tilde{f}} - u_f\|_{H_0^1(D)} \leq C\|f - \tilde{f}\|_{H^{-1}(D)}$$

G. Caloz and J. Rappaz, Handbook of Numerical Analysis, 1997.

- Similar results hold for Finite Element approximations and reduced order approximations.

Ensuring accurate predictions

- For very many PDEs we can bound the error between the solution u and the Finite Element approximation u_h as follows:

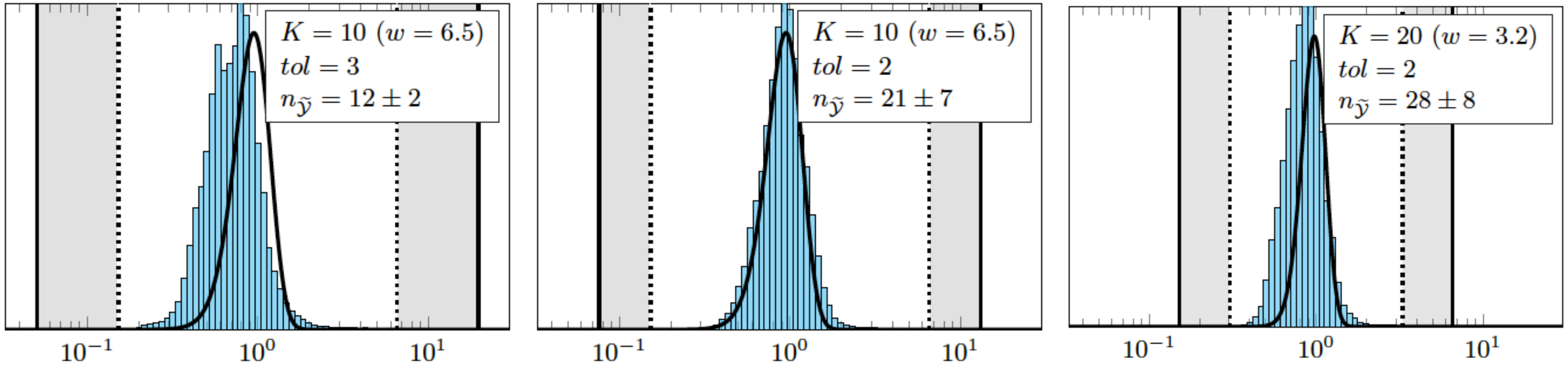
$$\|u - u_h\|_{H^1(D)} \leq Ch^k \|u\|_{H^{k+1}(D)} \quad \text{and} \quad \|u - u_h\|_{H^1(D)} \leq \Delta(u_h)$$

➡ Ensures convergence at a certain rate and allows us to assess accuracy of approximation.

- Similarly, we can bound error in quantity of interest and use bound to correct the quantity of interest.

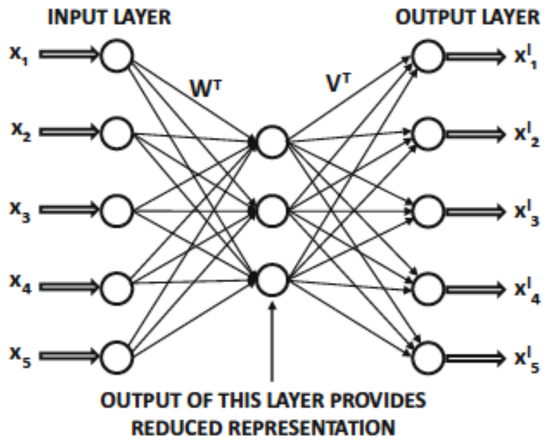
Probabilistic approaches for accuracy assessment

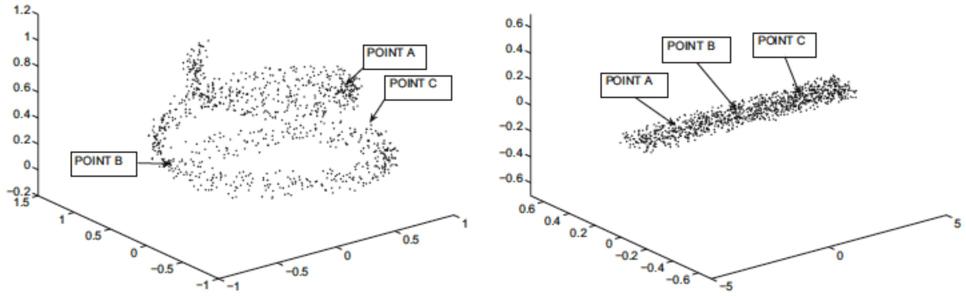
- Building statistical error models via Gaussian-process regression (M. Drohmann, K. Carlberg, SIAM J. Sci. Comput., 2015; S. Pagani, A. Manzoni, K. Carlberg, arXiv, 2019;...)
- Exploiting results from compressed sensing to build fast-to-evaluate unbiased estimator for error (Y. Cao, L. Petzold, SIAM J. Sci. Comput., 2004; K. Smetana, O. Zahm, A.T. Patera, SIAM J. Sci. Comput., 2019)
- Probabilistic Numerical Methods: Interpret standard numerical methods in a probabilistic manner; Numerical methods solve an inference task (P. Henning, M. A. Osborne, M. Girolami, Proc. R. Soc. A, 2015; Owhadi, MMS, 2015; Owhadi, SIAM Rev., 2017,...)



How to combine traditional methods and ML

- Stabilization of Neural Networks:
  - Interpret (simplified) Residual Network as discretization of ordinary differential equation ⟶ Derive stability criteria and develop stable networks (E. Haber and L. Ruthotto, Inverse Problems 17)
- Exploit connections between autoencoders and matrix decompositions:
  - Goal: Find matrix decomposition $A \approx UV^T$ such that $\| A - UV^T \|_F^2$ is minimal. That is realized by Singular Value decomposition but also by autoencoders with linear activation





(a) A nonlinear pattern in three dimensions (b) A reduced data set in two dimensions

C. C. Aggarwal, Neural Networks and Deep Learning, Springer, 2018.

How to combine traditional methods and ML

- Stabilization of Neural Networks:
  - Interpret (simplified) Residual Network as discretization of ordinary differential equation $\longrightarrow$ Derive stability criteria and develop stable networks (E. Haber and L. Ruthotto, Inverse Problems 17)
- Exploit connections between autoencoders and matrix decompositions:
  - Goal: Find matrix decomposition $A \approx UV^T$ such that $\| A - UV^T \|_F^2$ is minimal. That is realized by Singular Value decomposition but also by autoencoders with linear activation
- Physics-informed neural networks (M. Raissi, P. Perdikaris, G.E. Karniadakis, 17, 18, 19)
- Bayesian/probabilistic framework (e.g.: N. C. Nguyen et al, SIAM J. Sci. Comput, 2016)
- Data assimilation

Questions or comments?