Physics-informed Neural Networks for quasi-static Maxwell‘s Equations

Fraunhofer-Institut für Integrierte Systeme und Bauelemente-technologie IISB

Philipp Brendel, A. Rosskopf, V. Medvedev, 29.08.2022
Fraunhofer IISB Erlangen, Germany

Content

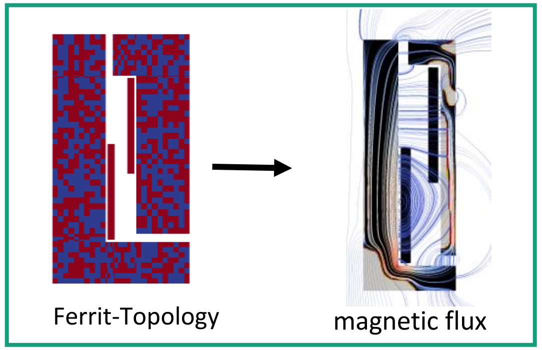
1. Introduction
2. Modeling
   - Quasi-static Maxwell Equation
   - Non-Dimensionalization
3. Implementation
   - Frameworks
   - Use-Cases & Reference Solutions
4. PINNs in Practice
   - Adaptive Loss-Weighting
   - Adaptive Sampling
5. Results & Outlook

Chapter 01

Introduction

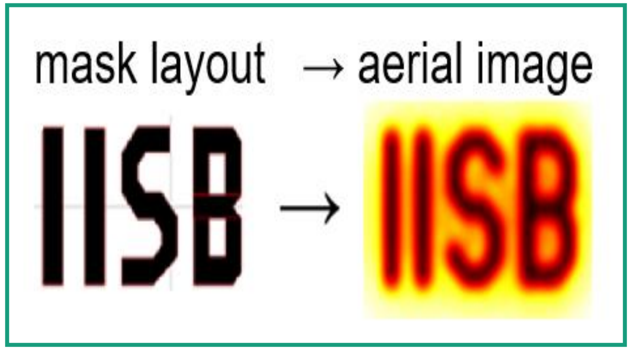
Motivation

- Applications in Power-Electronics require accurate simulation of the underlying Physics in order to optimize Devices or Processes
- accurate numerical Schemes and Softwares (FEM, FD, etc.) usually exist but can be:
  → computationally demanding
  → even more demanding at scale for optimization purposes with many degrees of freedom
- traditional ML-Approaches:
  → quick inference times on trained Models
  → limited possibility to include physics
  → „black-box“



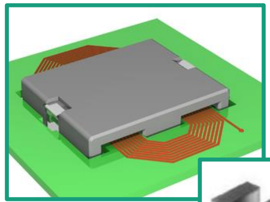
Optimization of Ferrit Cores inside Converter

(Reduction of Losses)



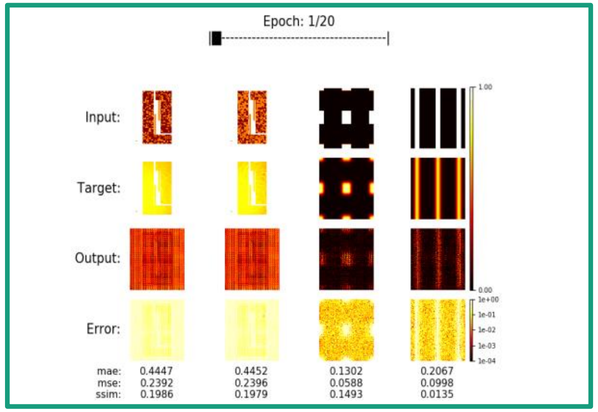
Optimization of Mask Layouts in EUV-Lithography

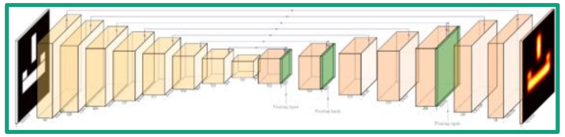
(Accurate Imaging)

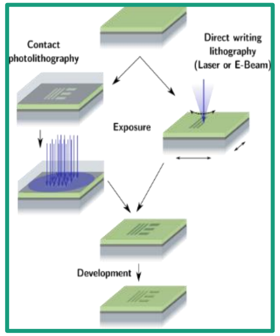


[1]

[2]







[3]

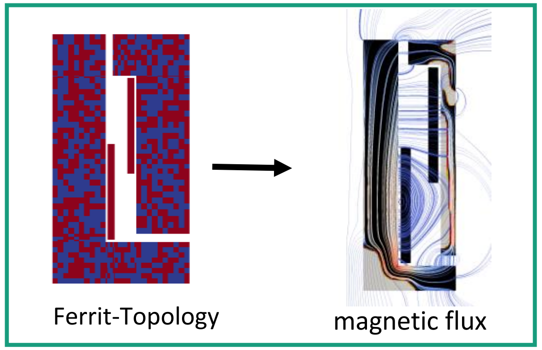
 29.08.2022

Fraunhofer IISB

Motivation

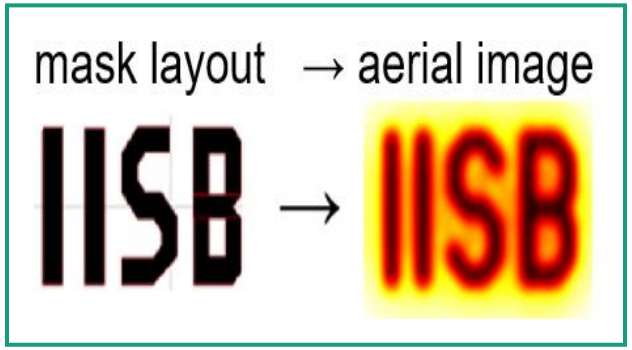
- Applications in Power-Electronics require accurate simulation of the underlying Physics in order to optimize Devices or Processes
- accurate numerical Schemes and Softwares (FEM, FD, etc.) usually exist but can be:
  → computationally demanding
  → even more demanding at scale for optimization purposes with many degrees of freedom
- traditional ML-Approaches:
  → quick inference times on trained Models
  → limited possibility to include physics
  → „black-box“

→ Physics-Informed Neural Networks



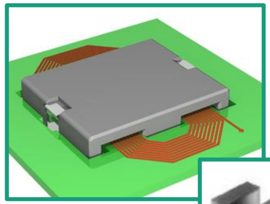
Optimization of Ferrit Cores inside Converter

(Reduction of Losses)



Optimization of Mask Layouts in EUV-Lithography

(Accurate Imaging)



[1]

[2]

$$\sigma i\omega\overline{A}_z(x,y) + \frac{1}{\mu}\left(-\frac{\partial^2\overline{A}_z(x,y)}{\partial x^2} - \frac{\partial^2\overline{A}_z(x,y)}{\partial y^2}\right) = \overline{J}_z(x,y) \qquad (23)$$

Randbedingungen:

$$A_z n_x = 0,\ A_z n_y = 0 \text{ auf } \Gamma$$

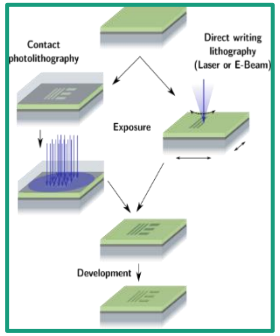
Interfacebedingungen:

$$(A_{z2} - A_{z1})n_y = 0,\ (A_{z2} - A_{z1})n_x = 0 \text{ auf } \Sigma_{12}$$

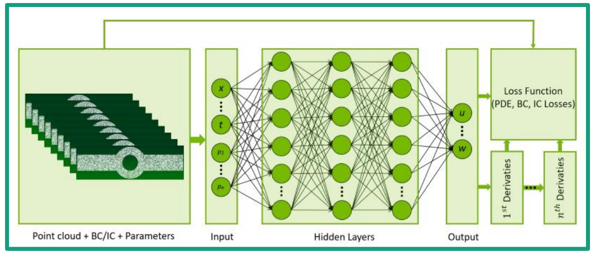
$$\left(\frac{1}{\mu_1}\frac{\partial A_{z1}}{\partial x} - \frac{1}{\mu_2}\frac{\partial A_{z2}}{\partial x}\right)n_z = 0,\ \left(\frac{1}{\mu_1}\frac{\partial A_{z1}}{\partial y} - \frac{1}{\mu_2}\frac{\partial A_{z2}}{\partial y}\right)n_z = 0,$$

$$\left(\frac{1}{\mu_2}\frac{\partial A_{z2}}{\partial x} - \frac{1}{\mu_1}\frac{\partial A_{z1}}{\partial x}\right)n_x + \left(\frac{1}{\mu_2}\frac{\partial A_{z2}}{\partial y} - \frac{1}{\mu_1}\frac{\partial A_{z1}}{\partial y}\right)n_y = 0 \text{ auf } \Sigma_{12}$$

$$\sigma_1 n_z\overline{A}_{z1} - \sigma_2 n_z\overline{A}_{z2} = 0 \text{ auf } \Sigma_{12}.$$



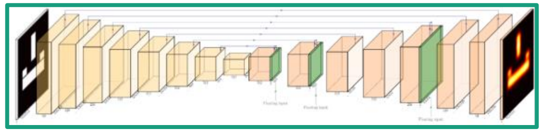
[3]

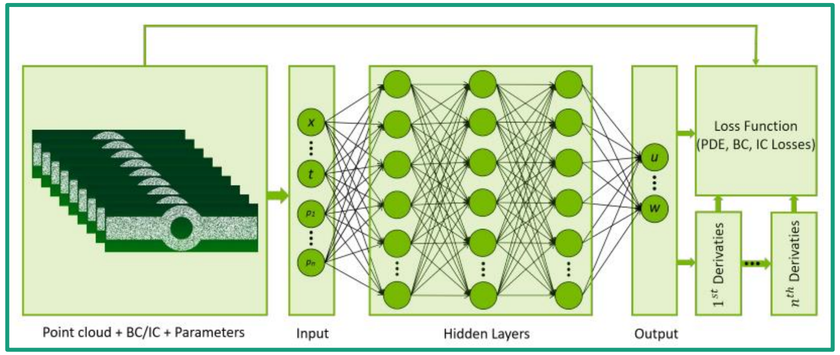


Physics-Informed Neural Networks

- notion first introduced by Raissi et al. in 2017 [4]

- Conceptual Difference in how Neural Networks are trained:
  - <u>classical Neural Network</u>:
    - „data-supervised Learning“
    - e.g.: Input-Image → Output-Image
    - Loss-Function: e.g. MSE over all pixels
  - <u>Physics-Informed Neural Network:</u>
    - „physics-supervised“ Learning
    - e.g.: (x, y, t) → u(x,y,t) → (du/dx, du/dy, du/dt, …)
    - Loss-Function: Residuals on PDE, BCs, ICs



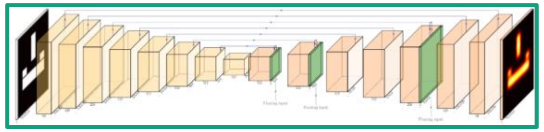


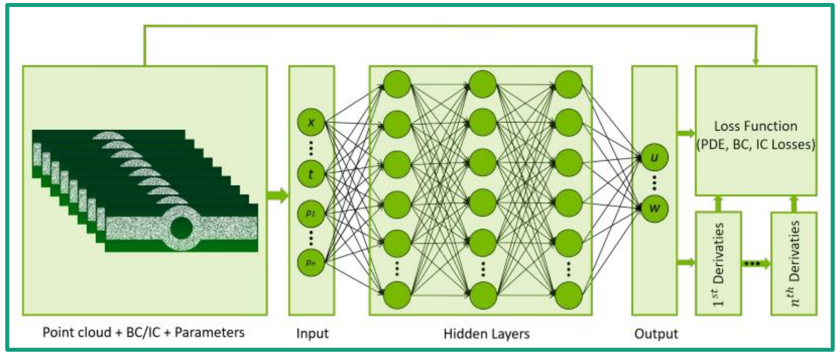
Physics-Informed Neural Networks

- notion first introduced by Raissi et al. in 2017 [4]

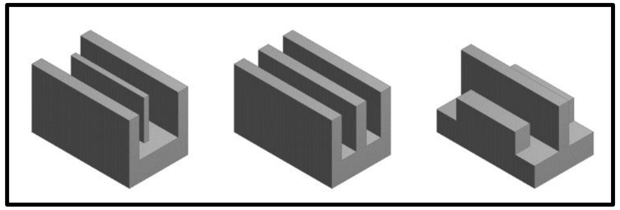
- Conceptual Difference in how Neural Networks are trained:
  - classical Neural Network:
    - „data-supervised Learning“
    - e.g.: Input-Image → Output-Image
    - Loss-Function: e.g. MSE over all pixels

  - Physics-Informed Neural Network:
    - „physics-supervised“ Learning
    - e.g.: (x, y, t) → u(x,y,t) → (du/dx, du/dy, du/dt, …)
    - Loss-Function: Residuals on PDE, BCs, ICs





- Benefits
  - no data required → explicit supervision by Physics
  - implicit meshing → adaptive strategies can be applied
  - generalization → additional Inputs e.g. for different Geometries

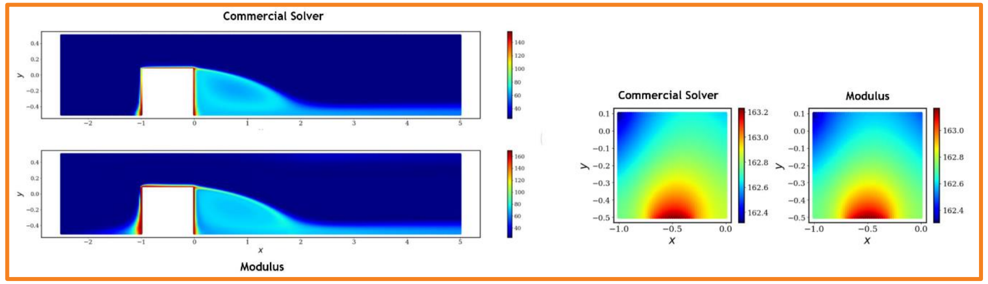


Parametrization of Heat Sink Geometries [6]

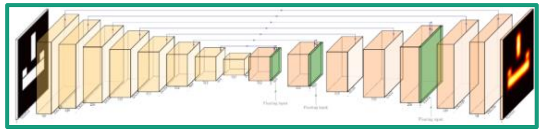
 29.08.2022

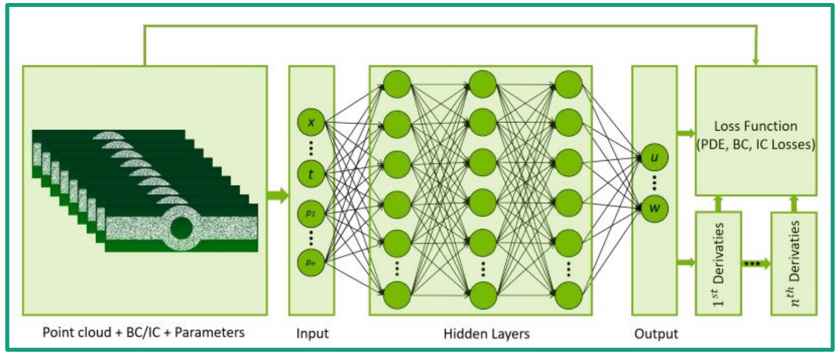
Physics-Informed Neural Networks

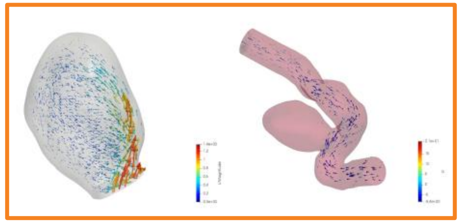
- notion first introduced by Raissi et al. in 2017 [4]
- Conceptual Difference in how Neural Networks are trained:
  - classical Neural Network:
    - „data-supervised Learning“
    - e.g.: Input-Image → Output-Image
    - Loss-Function: e.g. MSE over all pixels
  - Physics-Informed Neural Network:
    - „physics-supervised“ Learning
    - e.g.: (x, y, t) → u(x,y,t) → (du/dx, du/dy, du/dt, …)
    - Loss-Function: Residuals on PDE, BCs, ICs
- Benefits
  - no data required → explicit supervision by Physics
  - implicit meshing → adaptive strategies can be applied
  - generalization → additional Inputs e.g. for different Geometries
- Increasing Amounts of Publications and (Open-Source) Software [5]



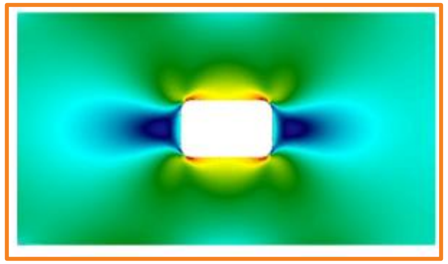
Temperature Simulation in Fluids & Solids [7]



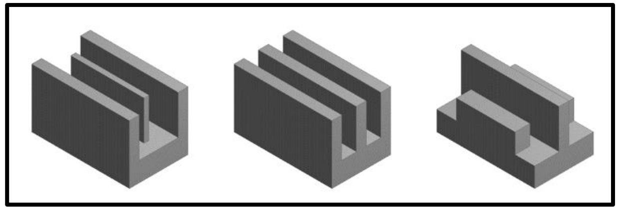




Blood flow simulation in an Aneurysm [7]



Simulation of aircraft cabin panel stress [7]

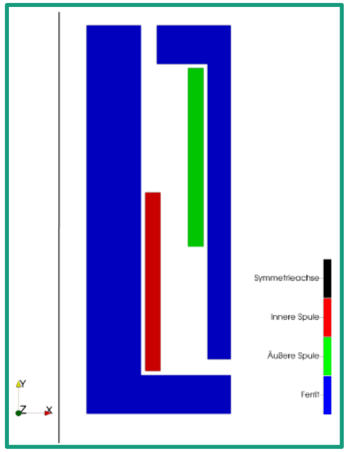


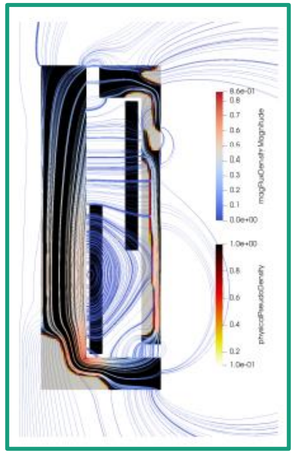
Parametrization of Heat Sink Geometries [6]

Fraunhofer IISB

A Use-Case from Power-Electronics

Topology-Optimization of a Power Converter





Goals

1. Approximation of magnetic Fields by solving quasi-static Maxwell‘s Equation
2. Calculation of relevant Quantities like Inductivities and Coupling *k*
3. Optimization of Topology with respect to those Quantities

$$M := L_{12} = L_{21}.$$

$$k = \frac{M}{\sqrt{L_{11}L_{22}}}.$$

$$L_{21} = \frac{\Phi_{12}}{i_1}$$

Fraunhofer IISB

Chapter 02

Modeling

Modeling

Quasi-static Maxwell Equation

- K.Angermeier: „Topology-Optimization of inductive Components“ [8]

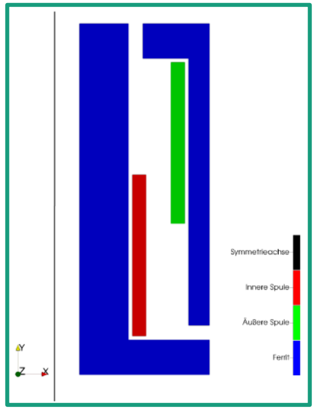
$$\nabla \times \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t} \qquad (1)$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \qquad (2)$$

$$\nabla \cdot \boldsymbol{B} = 0 \qquad (3)$$

$$\nabla \cdot \boldsymbol{D} = \rho \qquad (4)$$

$$\begin{aligned} \boldsymbol{B} &= \mu \boldsymbol{H} \\ \boldsymbol{J} &= \sigma \boldsymbol{E} \qquad (5) \\ \boldsymbol{D} &= \varepsilon \boldsymbol{E}. \end{aligned}$$



Fraunhofer IISB

Modeling

Quasi-static Maxwell Equation

- K.Angermeier: „Topology-Optimization of inductive Components“ [8]

$$\nabla \times \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t} \qquad (1)$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \qquad (2)$$

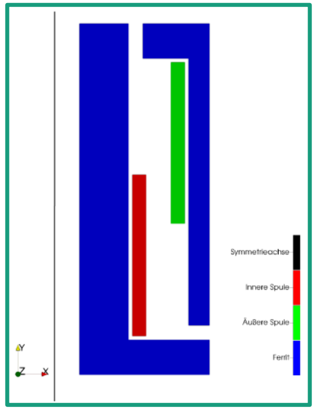
$$\nabla \cdot \boldsymbol{B} = 0 \qquad (3)$$

$$\nabla \cdot \boldsymbol{D} = \rho \qquad (4)$$

$$\boldsymbol{B} = \mu \boldsymbol{H}$$

$$\boldsymbol{J} = \sigma \boldsymbol{E} \qquad (5)$$

$$\boldsymbol{D} = \varepsilon \boldsymbol{E}.$$



- no free electric charges/currents
- Existence of magnetic Vector Potential A, s.t.

$$\boxed{\boldsymbol{B} = \nabla \times \boldsymbol{A}}$$

- Coulomb Gauging for Uniqueness of A

$$\boxed{\nabla \cdot \boldsymbol{A} = 0}$$

- rotational symmetry, current J only in z-direction

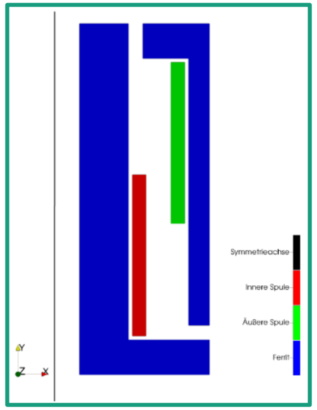
$$\Rightarrow \quad \boxed{\sigma \frac{\partial \boldsymbol{A}}{\partial t} + \nabla \times \frac{1}{\mu} \nabla \times \boldsymbol{A} = \boldsymbol{J}}$$

Modeling

Quasi-static Maxwell Equation

- K.Angermeier: „Topology-Optimization of inductive Components“ [8]

$$
\begin{aligned}
\nabla \times \boldsymbol{H} &= \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t} && (1)\\
\nabla \times \boldsymbol{E} &= -\frac{\partial \boldsymbol{B}}{\partial t} && (2)\\
\nabla \cdot \boldsymbol{B} &= 0 && (3)\\
\nabla \cdot \boldsymbol{D} &= \rho && (4)\\
\boldsymbol{B} &= \mu \boldsymbol{H} &&\\
\boldsymbol{J} &= \sigma \boldsymbol{E} && (5)\\
\boldsymbol{D} &= \varepsilon \boldsymbol{E}. &&
\end{aligned}
$$



- no free electric charges/currents
- Existence of magnetic Vector Potential A, s.t.

$$\boxed{\boldsymbol{B} = \nabla \times \boldsymbol{A}}$$

- Coulomb Gauging for Uniqueness of A

$$\boxed{\nabla \cdot \boldsymbol{A} = 0}$$

- rotational symmetry, current J only in z-direction

$$\boxed{\sigma \frac{\partial \boldsymbol{A}}{\partial t} + \nabla \times \frac{1}{\mu} \nabla \times \boldsymbol{A} = \boldsymbol{J}}$$

- time-dependence on J and A

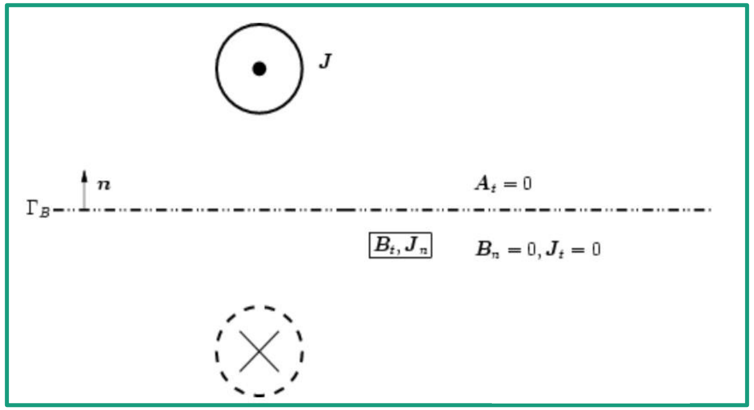
$$J_z(x, y; t) = e^{i\omega t} \overline{J}_z(x, y)$$

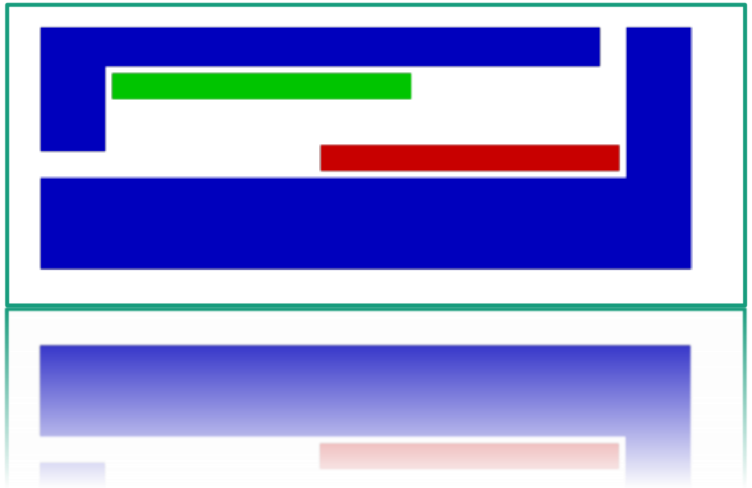
$$A_z(x, y; t) = e^{i\omega t} \overline{A}_z(x, y)$$

$$\boxed{\sigma i \omega \overline{A}_z(x, y) + \frac{1}{\mu} \left(-\frac{\partial^2 \overline{A}_z(x, y)}{\partial x^2} - \frac{\partial^2 \overline{A}_z(x, y)}{\partial y^2} \right) = \overline{J}_z(x, y)}$$

Modeling

Boundary Conditions





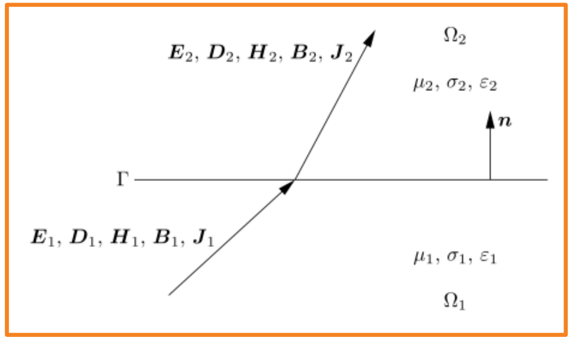
- anti-symmetric Mirroring of source currents by enforcing zero magnetic flux across Boundary, cf. Harlander, 2002 [9]:

$$\boxed{\boldsymbol{n} \cdot \boldsymbol{B} = 0}$$

- Biro, Preis, 1989 [10]: equivalence to homogeneous dirichlet conditions on magnetic Vector Potential A

$$\boxed{A_z n_x = 0,\ A_z n_y = 0}$$

Modeling
Interface Conditions



$$H(\mathrm{div};\Omega) := \left\{ v \in L^2(\Omega)^3 : \nabla \cdot v \in L^2(\Omega) \right\}$$
$$H(\mathrm{rot};\Omega) := \left\{ v \in L^2(\Omega)^3 : \nabla \times v \in L^2(\Omega)^3 \right\}$$

$$\boldsymbol{H}(\cdot,t) \in H(\mathrm{rot};\Omega),\ \boldsymbol{E}(\cdot,t) \in H(\mathrm{rot};\Omega)$$
$$\boldsymbol{B}(\cdot,t) \in H(\mathrm{div};\Omega),\ \boldsymbol{J}(\cdot,t) \in H(\mathrm{div};\Omega)$$

- For L2-Regularity, we require continuity across interfaces [11]:

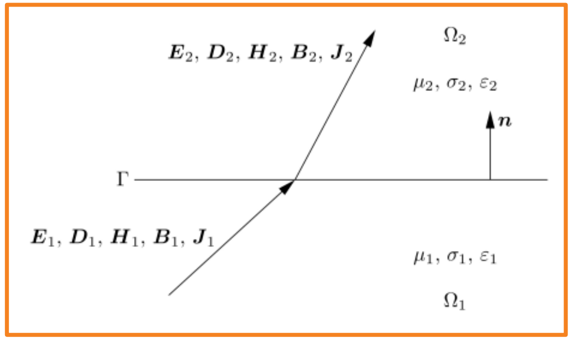
$$\boxed{\boldsymbol{B} = \nabla \times \boldsymbol{A}}$$

  - normal component of magnetic flux → $B_1 \cdot n = B_2 \cdot n$

$$\boxed{(A_{z2} - A_{z1})n_y = 0,\ (A_{z2} - A_{z1})n_x = 0}$$

Modeling

Interface Conditions



$$\boxed{\begin{aligned} H(\mathrm{div};\Omega) &:= \left\{v \in L^2(\Omega)^3 : \nabla \cdot v \in L^2(\Omega)\right\} \\ H(\mathrm{rot};\Omega) &:= \left\{v \in L^2(\Omega)^3 : \nabla \times v \in L^2(\Omega)^3\right\} \end{aligned}}$$

$$\boxed{\begin{aligned} \boldsymbol{H}(\cdot,t) \in H(\mathrm{rot};\Omega),\ \boldsymbol{E}(\cdot,t) \in H(\mathrm{rot};\Omega) \\ \boldsymbol{B}(\cdot,t) \in H(\mathrm{div};\Omega),\ \boldsymbol{J}(\cdot,t) \in H(\mathrm{div};\Omega) \end{aligned}}$$

- For L2-Regularity, we require continuity across interfaces [11]:
  - normal component of magnetic flux → $B_1 \cdot n = B_2 \cdot n$

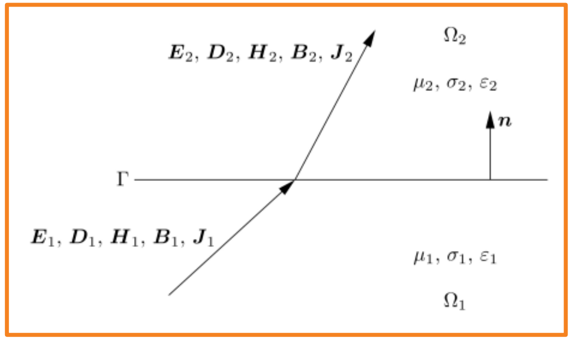
$$\boxed{\boldsymbol{B} = \nabla \times \boldsymbol{A}} \Rightarrow \boxed{(A_{z2} - A_{z1})n_y = 0,\ (A_{z2} - A_{z1})n_x = 0}$$

  - normal component of current → $J_1 \cdot n = J_2 \cdot n$

$$\boxed{\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}},\ \boxed{\boldsymbol{J} = \sigma \boldsymbol{E}} \Rightarrow \boxed{\sigma_1 n_z \overline{A}_{z1} - \sigma_2 n_z \overline{A}_{z2} = 0}$$

Modeling

Interface Conditions



$$H(\mathrm{div};\Omega) := \left\{v \in L^2(\Omega)^3 : \nabla \cdot v \in L^2(\Omega)\right\}$$
$$H(\mathrm{rot};\Omega) := \left\{v \in L^2(\Omega)^3 : \nabla \times v \in L^2(\Omega)^3\right\}$$

$$\boldsymbol{H}(\cdot,t) \in H(\mathrm{rot};\Omega),\ \boldsymbol{E}(\cdot,t) \in H(\mathrm{rot};\Omega)$$
$$\boldsymbol{B}(\cdot,t) \in H(\mathrm{div};\Omega),\ \boldsymbol{J}(\cdot,t) \in H(\mathrm{div};\Omega)$$

- For L2-Regularity, we require continuity across interfaces [11]:

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}$$

  - normal component of magnetic flux → $B_1 \cdot n = B_2 \cdot n$

$$(A_{z2} - A_{z1})n_y = 0,\ (A_{z2} - A_{z1})n_x = 0$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}$$

  - normal component of current → $J_1 \cdot n = J_2 \cdot n$

$$\boldsymbol{J} = \sigma \boldsymbol{E}$$

$$\sigma_1 n_z \overline{A}_{z1} - \sigma_2 n_z \overline{A}_{z2} = 0$$

  - tangential component of magnetic field intensity → $H_1 \cdot t = H_2 \cdot t$

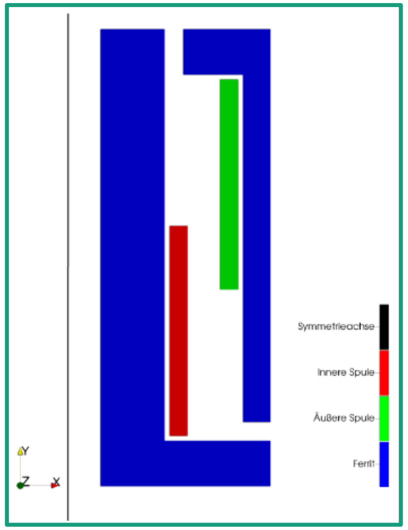
$$\boldsymbol{B} = \mu \boldsymbol{H}$$

$$\left(\frac{1}{\mu_1}\frac{\partial A_{z1}}{\partial x} - \frac{1}{\mu_2}\frac{\partial A_{z2}}{\partial x}\right) n_z = 0,\ \left(\frac{1}{\mu_1}\frac{\partial A_{z1}}{\partial y} - \frac{1}{\mu_2}\frac{\partial A_{z2}}{\partial y}\right) n_z = 0,$$
$$\left(\frac{1}{\mu_2}\frac{\partial A_{z2}}{\partial x} - \frac{1}{\mu_1}\frac{\partial A_{z1}}{\partial x}\right) n_x + \left(\frac{1}{\mu_2}\frac{\partial A_{z2}}{\partial y} - \frac{1}{\mu_1}\frac{\partial A_{z1}}{\partial y}\right) n_y = 0$$

Fraunhofer IISB

Modeling

Summary



Quasi-static Maxwell-Equation

$$\sigma i\omega \overline{A}_z(x,y) + \frac{1}{\mu}\left(-\frac{\partial^2 \overline{A}_z(x,y)}{\partial x^2} - \frac{\partial^2 \overline{A}_z(x,y)}{\partial y^2}\right) = \overline{J}_z(x,y)$$

$$\text{Randbedingungen:}$$

$$A_z n_x = 0,\ A_z n_y = 0 \text{ auf } \Gamma$$

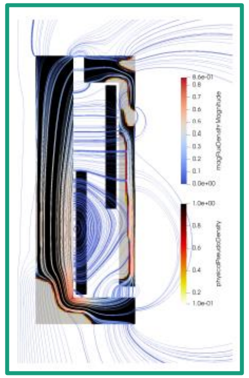
$$\text{Interfacebedingungen:}$$

$$(A_{z2} - A_{z1})n_y = 0,\ (A_{z2} - A_{z1})n_x = 0 \text{ auf } \Sigma_{12}$$

$$\left(\frac{1}{\mu_1}\frac{\partial A_{z1}}{\partial x} - \frac{1}{\mu_2}\frac{\partial A_{z2}}{\partial x}\right)n_z = 0,\ \left(\frac{1}{\mu_1}\frac{\partial A_{z1}}{\partial y} - \frac{1}{\mu_2}\frac{\partial A_{z2}}{\partial y}\right)n_z = 0,$$

$$\left(\frac{1}{\mu_2}\frac{\partial A_{z2}}{\partial x} - \frac{1}{\mu_1}\frac{\partial A_{z1}}{\partial x}\right)n_x + \left(\frac{1}{\mu_2}\frac{\partial A_{z2}}{\partial y} - \frac{1}{\mu_1}\frac{\partial A_{z1}}{\partial y}\right)n_y = 0 \text{ auf } \Sigma_{12}$$

$$\sigma_1 n_z \overline{A}_{z1} - \sigma_2 n_z \overline{A}_{z2} = 0 \text{ auf } \Sigma_{12}.$$

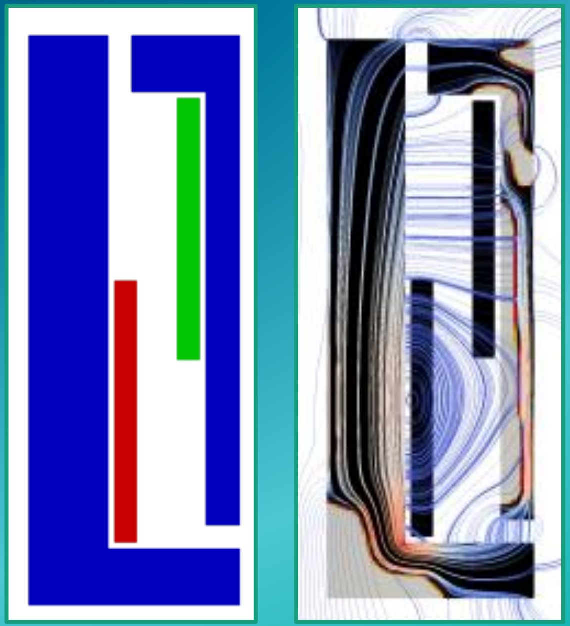


$$\boldsymbol{B} = \nabla \times A_z \boldsymbol{e_z} = \begin{pmatrix} \frac{\partial A_z}{\partial y} \\ \frac{\partial A_z}{\partial x} \\ 0 \end{pmatrix} = \begin{pmatrix} B_x \\ B_y \\ 0 \end{pmatrix}$$

 29.08.2022

Fraunhofer IISB

Modeling
Main Challenges



1. General Scaling and Stability of PDE
2. Nonlinear Material Properties at Interfaces between Air & Ferrit
3. Parametrization of Topology
4. Generalizable PINN-Training including geometry-specific Parameters

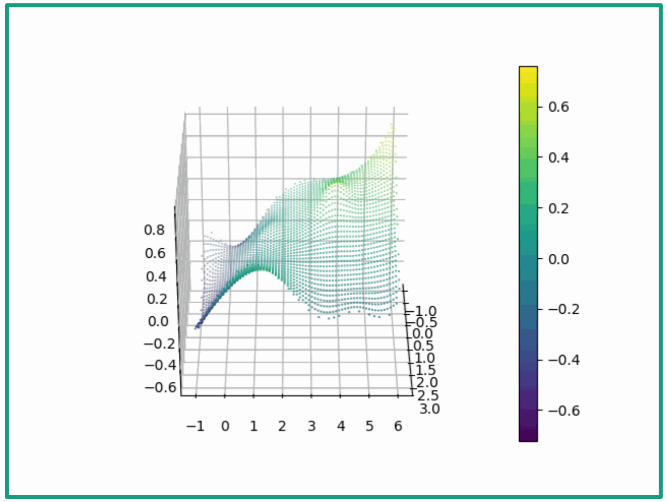
Fraunhofer IISB

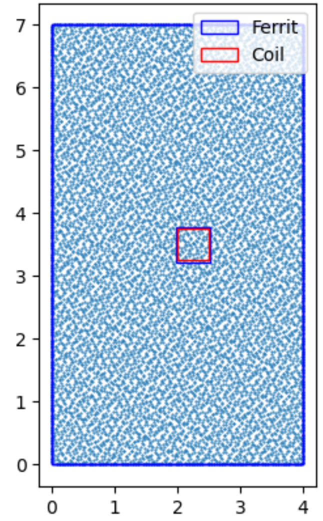
Modeling
Non-Dimensionalization

- Main Issues (presumably):
  - numerically small material parameters
  - permeability jumps by a factor of 2000 across interfaces

Property	Ferrit	Air
Permeability μ [H/m]	2.704e-3	1.255e-6
Conductivity σ [1/(Ω*m)]	5.8e-7	0.0

- Beltran-Pulido et. al, 2022 [12]
  - „… help to alleviate the effects of sharp changes of material properties across interfaces between materials"





Modeling

Non-Dimensionalization

- introducing rescaled, non-dimensionalized variables and parameters:

$$\boxed{\sigma i\omega A + \nabla(\frac{1}{\mu}(-\nabla A)) = J}$$

$$\overline{x} = \frac{x}{x^*}$$
$$\overline{A} = \frac{A}{A^*}$$
$$\overline{J} = \frac{J}{J^*}$$
$$\overline{\mu} = \frac{\mu}{\mu^*}$$
$$\overline{\sigma} = \frac{\sigma}{\sigma^*}$$

$$\mu^* = \frac{A^*}{(x^*)^2 J^*}$$
$$\sigma^* = \frac{J^*}{i\omega A^*}$$

$$\boxed{\overline{\sigma}\overline{A} + \overline{\nabla}(\frac{1}{\overline{\mu}}(-\overline{\nabla}\,\overline{A})) = \overline{J}}$$

Fraunhofer IISB

Modeling

Non-Dimensionalization

- introducing rescaled, non-dimensionalized variables and parameters:

$$\sigma i\omega A + \nabla\left(\frac{1}{\mu}(-\nabla A)\right) = J$$

$$\overline{x} = \frac{x}{x^*}, \quad \overline{A} = \frac{A}{A^*}, \quad \overline{J} = \frac{J}{J^*}, \quad \overline{\mu} = \frac{\mu}{\mu^*}, \quad \overline{\sigma} = \frac{\sigma}{\sigma^*}$$

$$\mu^* = \frac{A^*}{(x^*)^2 J^*}, \quad \sigma^* = \frac{J^*}{i\omega A^*}$$

$$\overline{\sigma}\,\overline{A} + \overline{\nabla}\left(\frac{1}{\overline{\mu}}(-\overline{\nabla}\,\overline{A})\right) = \overline{J}$$

- choice of constants, cf. [12], s.t.:

$$\frac{1}{\overline{\mu}} \in [0.05, 100]$$

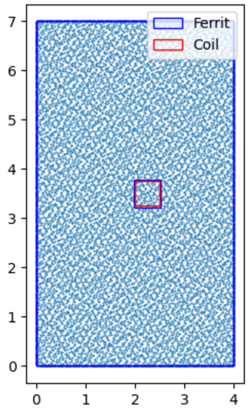
$$x^* = 7.0\ \mathrm{cm}, \quad A^* = 4.9\ \frac{\mathrm{mWb}}{\mathrm{m}}, \quad \mu^* = 1.352 \cdot 10^{-4}\ \frac{\mathrm{H}}{\mathrm{m}}$$

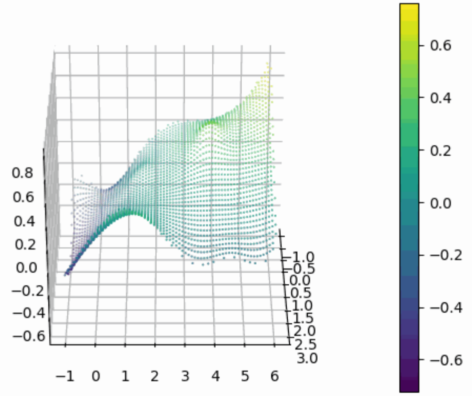
$$J^* = 1.352 \cdot 10^4\ \frac{\mathrm{A}}{\mathrm{m}^2}, \quad \sigma^* = 0.2759 \cdot 10^4$$

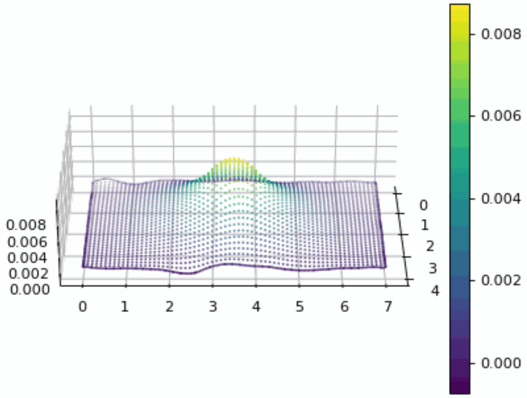
Modeling

Non-Dimensionalization

- (so far only) empirical evidence that even really simple use-cases behave a lot more stable in the non-dimensionalized setting







Chapter 03

Implementation

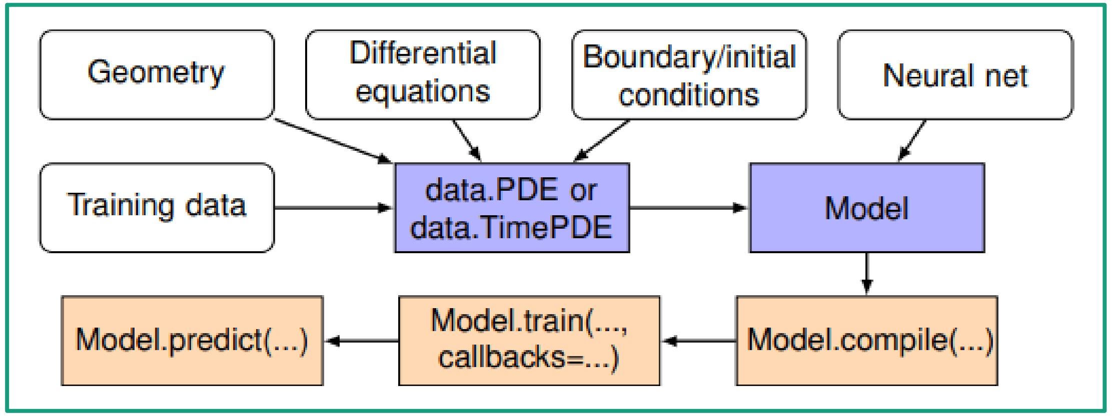
Implementation
PINN-Frameworks

- technical simplicity of the approach:
  - many open-source implementations available:
    - stand-alone Code for highly specific Problems
    - more general Frameworks with active development and maintenance

- Nvidia Modulus [6]:
  - formerly known as *SimNet™* [13]
  - more advanced, professionally maintained
  - more demanding/difficult to setup, especially in HPC-Environments

- DeepXDE [14]:
  - Python-based Framework
  - support for all common Machine-Learning Libraries (TensorFlow, PyTorch, …)
  - active online-community with many Publications connected to it

→ currently Framework of Choice for our work:
→ introduction of several Use-Cases to reflect different Complexity-Aspects

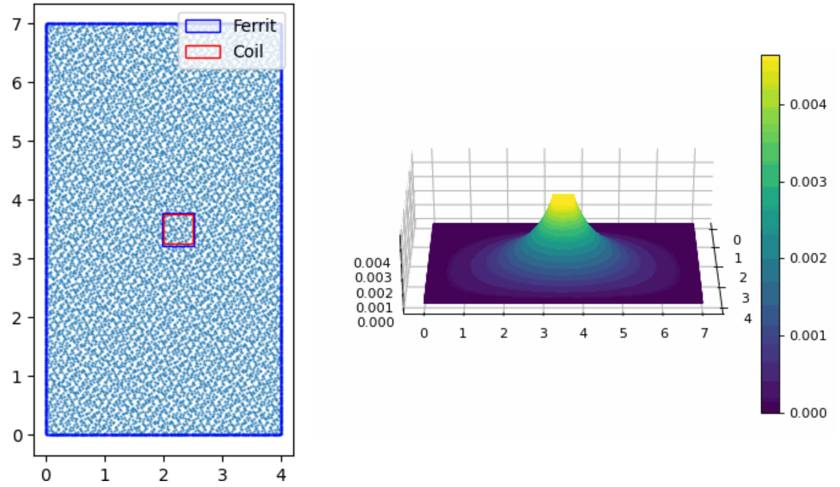


DeepXDE-Flowchart[14]

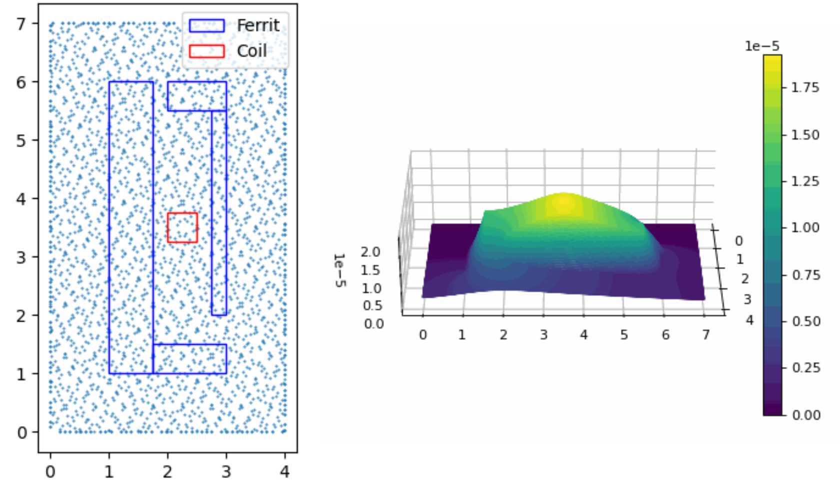
Implementation

Use-Cases & Reference Solutions

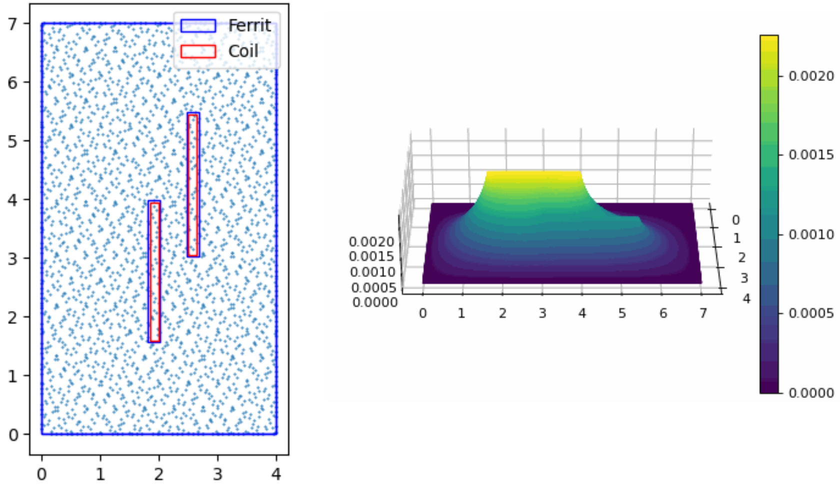
- a1:
  - single coil
  - surrounded by ferrit



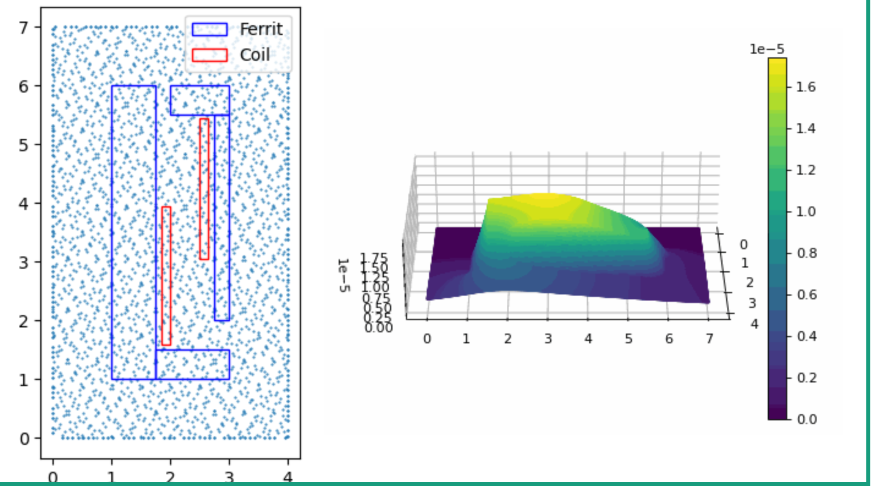
- a2:
  - single coil
  - surronded by air/ferrit



- b1:
  - two coils
  - surrounded by ferrit



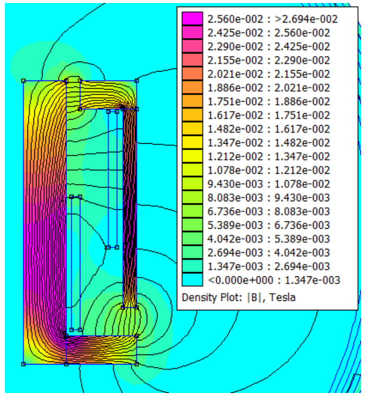
- b2:
  - two coils
  - surrounded by air/ferrit



Implementation

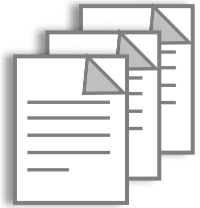
Reference Solutions

FEMM-Simulation

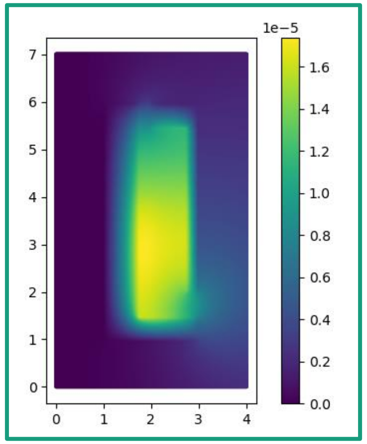


LUA

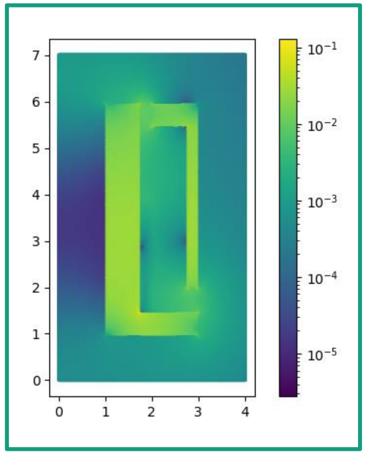
CSV-Format



magnetic Vector Potential A



magnetic Field Intensity B



- open-source Software FEMM: Finite-Element-Method Magnetics [15]

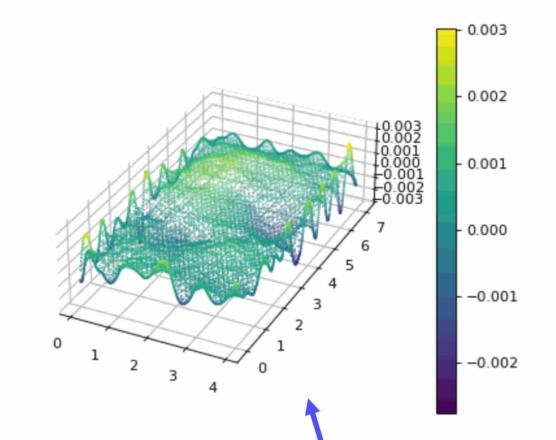
Implementation

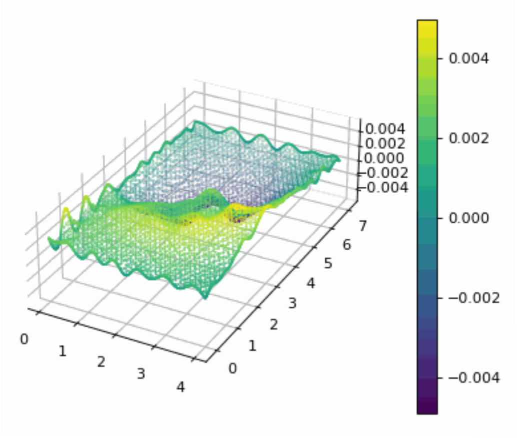
Reference Solutions

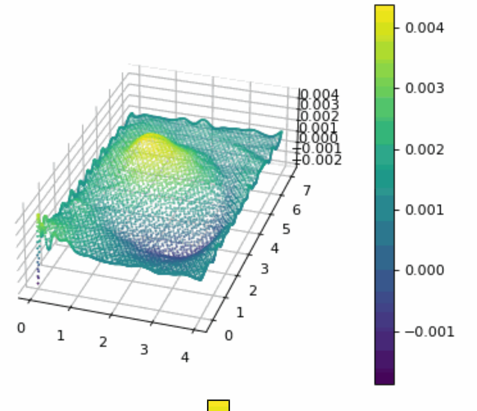
Need for Reference Solutions in Practice

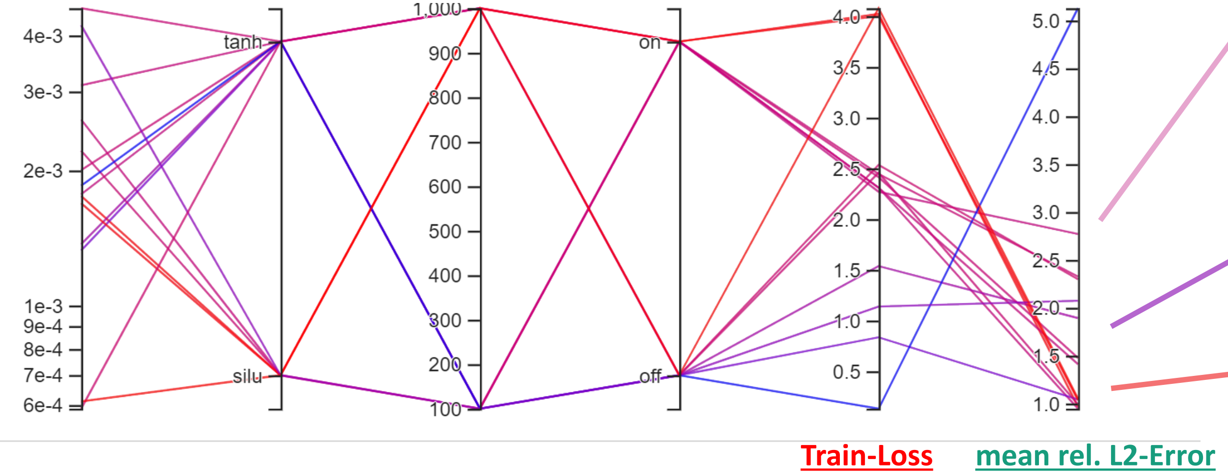
Losses:

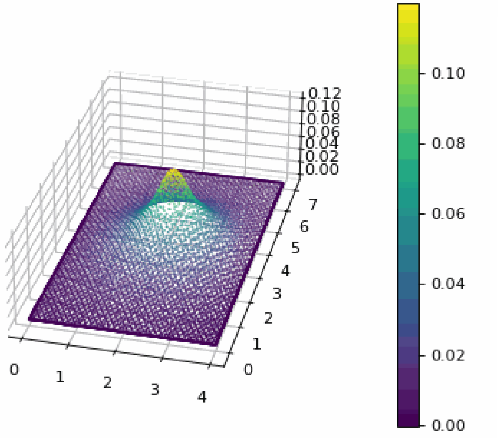
- numerically depend on weights
- might be misbalanced and misleading
- evaluation via relative L2-Error more helpful, e.g. to determine good Hyperparameters











 29.08.2022

Fraunhofer IISB

Chapter 04

PINNs in Practice

PINNs in Practice

(Adaptive) Loss-Weighting

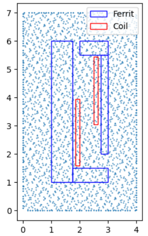
- Different Types and Numbers of Loss-Terms depending on Problem Formulation
  - PDE-Loss(es)
    - Loss for PDE-Residual
    - Loss for B = rot(A)

$$\boxed{\sigma i\omega A + \nabla\left(\frac{1}{\mu}(-\nabla A)\right) = J}$$

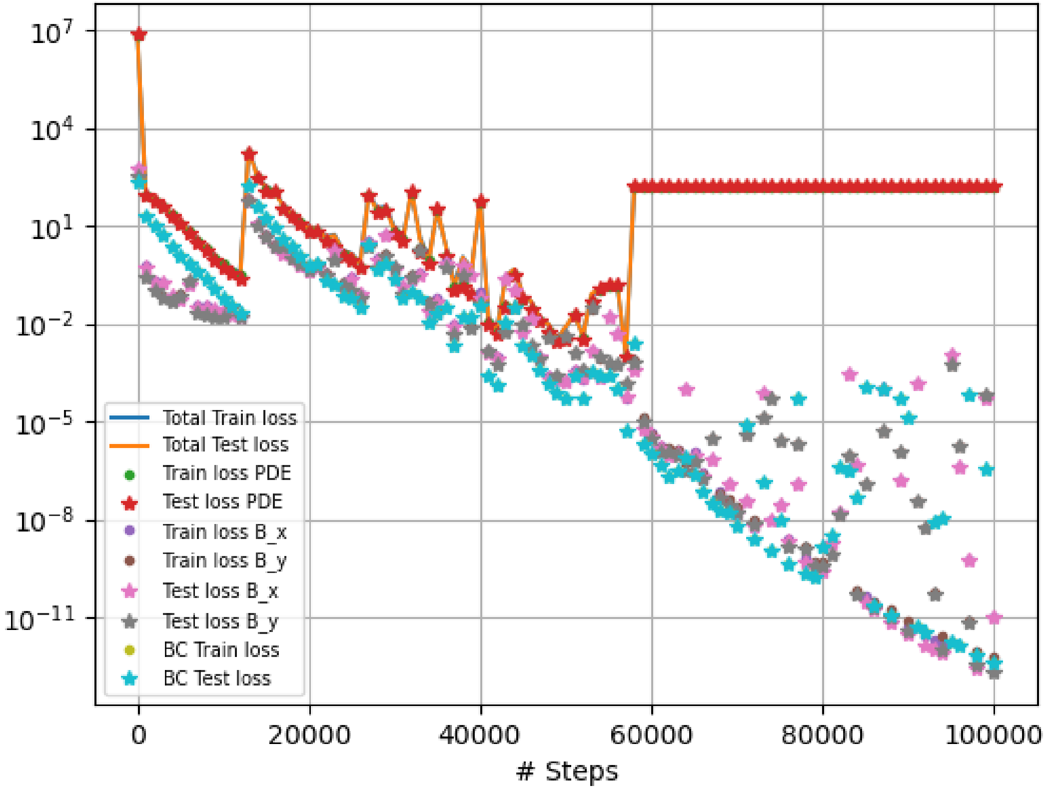
$$\boxed{\boldsymbol{B} = \nabla \times A_z \boldsymbol{e_z} = \begin{pmatrix} \frac{\partial A_z}{\partial y} \\ \frac{\partial A_z}{\partial x} \\ 0 \end{pmatrix} = \begin{pmatrix} B_x \\ B_y \\ 0 \end{pmatrix}}$$

  - Boundary Condition Losses
    - Dirichlet-Type Conditions can be enforced directly via Network-Structure
    - all other Types lead to additional Loss-Terms

- Interface Condition Losses
  - additional Terms for any Interfaces
  - amount increasing with complexity of topology



- Mathematically Advanced Strategies exist based on:
  - Neural Tangent Kernel [16] → Spectral Bias of Neural Networks
  - augmented Lagrangian Method [17] → adaptive Imitation of Lagrange-Multipliers



PINNs in Practice

(Adaptive) Loss-Weighting

Example: Weighting of Interface Condition Loss

$$\boxed{(A_{z2} - A_{z1})n_y = 0,\ (A_{z2} - A_{z1})n_x = 0 \text{ auf } \Sigma_{12}}$$

Interface Condition-Loss **non-zero at very few training points**

→ very small averaged loss per default

→ no incentive to reduce the IC-Loss ahead of other loss-terms

- loss weights: [1, 1, 1, 1, 1, 1, 100]

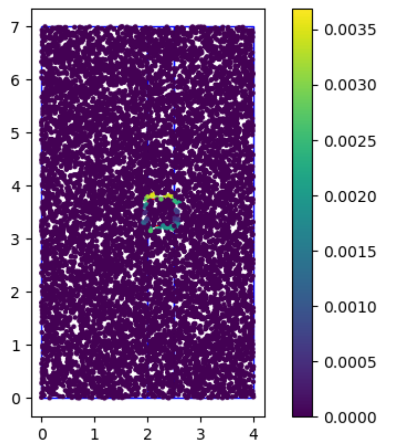
```
Step      Train loss
0         [1.05e-05, 1.26e+00, 1.23e-01, 4.91e-06, 0.00e+00, 5.82e-02, 7.47e-02]
1000      [3.02e-09, 2.19e-06, 2.27e-06, 1.98e-16, 0.00e+00, 8.78e-12, 8.64e-07]
```

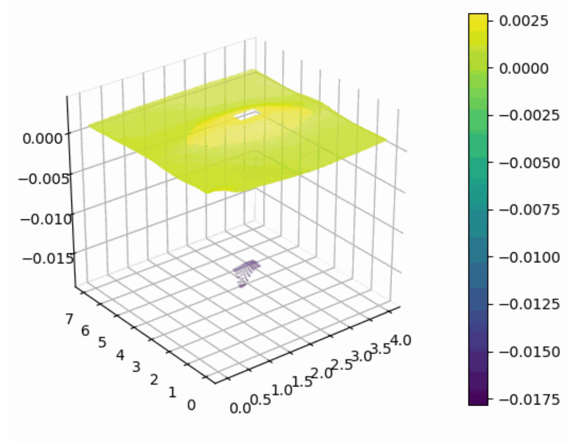
- loss weights: [1, 1, 1, 100000, 1, 1, 100]

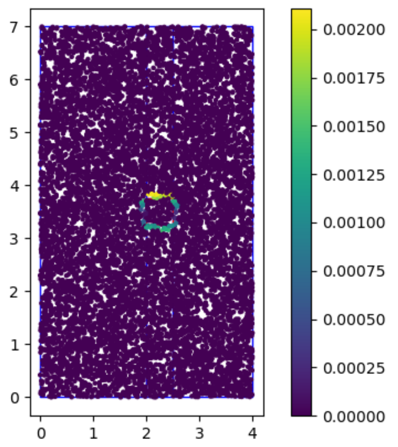
```
Step      Train loss
0         [1.05e-05, 1.26e+00, 1.23e-01, 4.91e-01, 0.00e+00, 5.82e-02, 7.47e-02]
1000      [6.95e-07, 2.24e-06, 6.28e-06, 1.00e-13, 0.00e+00, 3.75e-13, 3.93e-05]
```

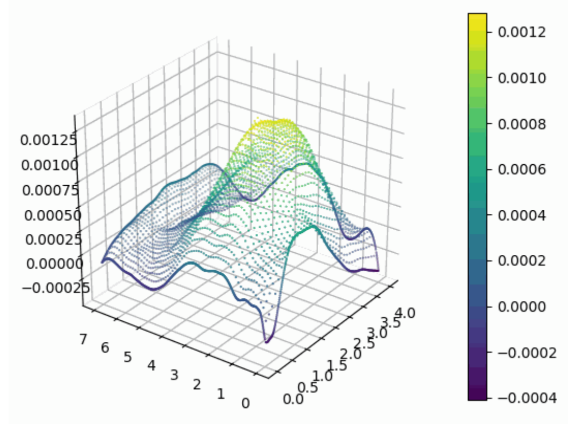
IC-Residual

Output





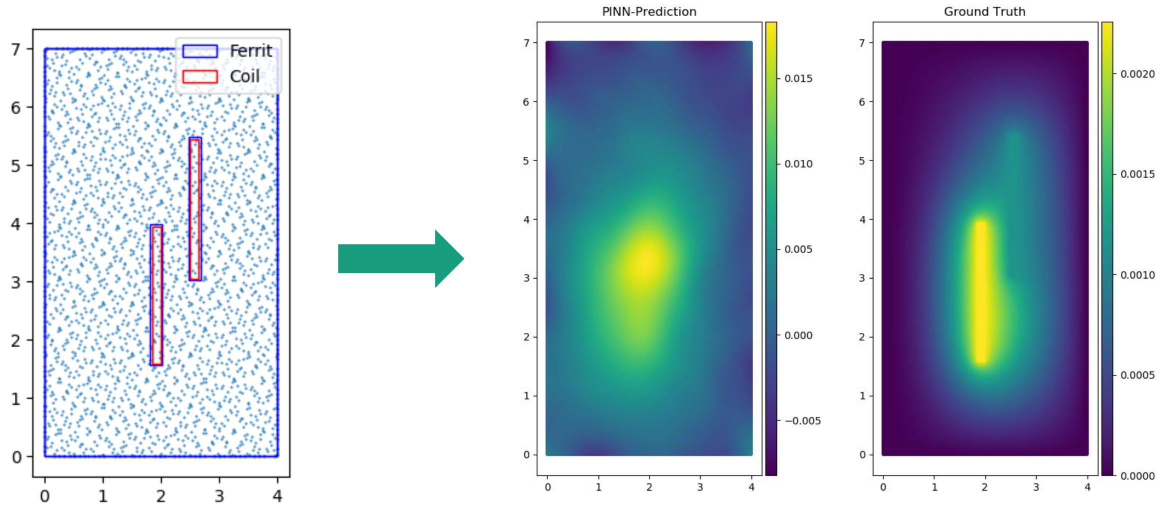




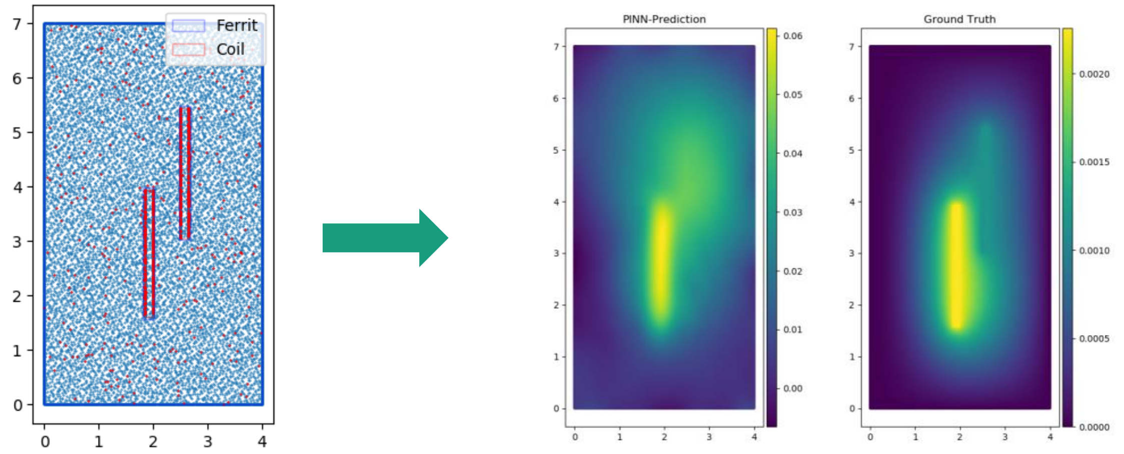
PINNs in Practice

Adaptive Sampling

- uniform sampling does not resolve the most dynamic areas sufficiently:
  - interfaces where material properties change
  - conductive parts of the material



- Residual-based adaptive Refinement (RAR)
  - based on the highest Residuals (PDE, BC, IC,..) new points in the domain are selected for Training
  - leads to new Hyperparameters
    - how many points to add?
    - at which point during training should this be done?
    - how often and how long (# epochs) should a Refinement-Step be done?



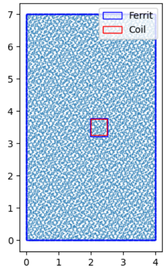
Fraunhofer IISB

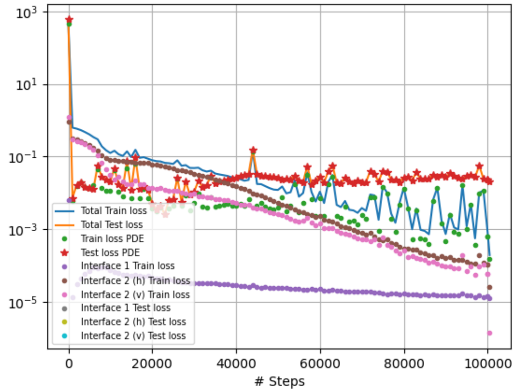
Chapter 05

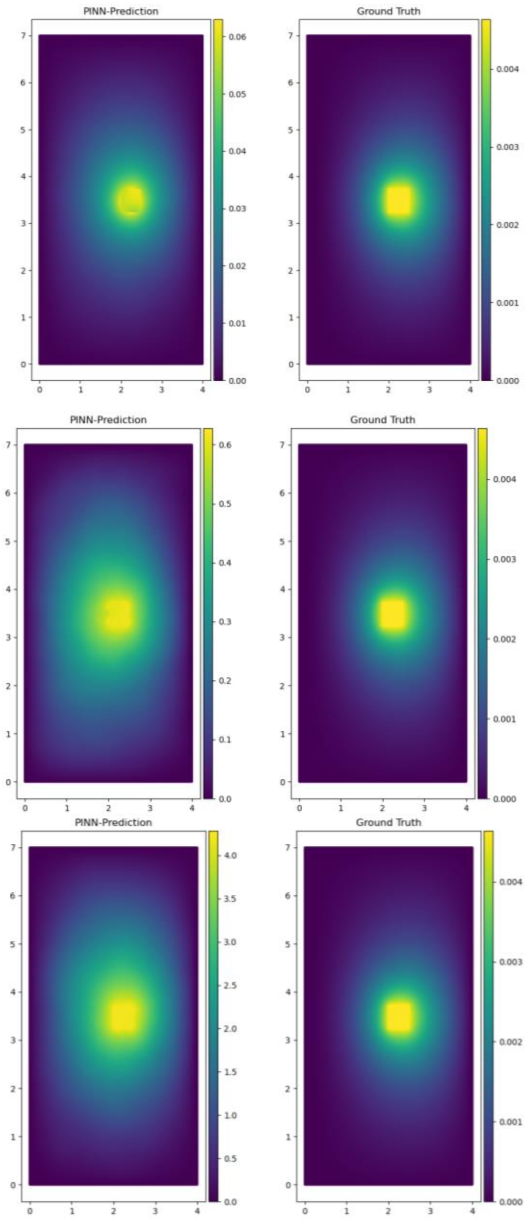
Results & Outlook

Intermediate Results

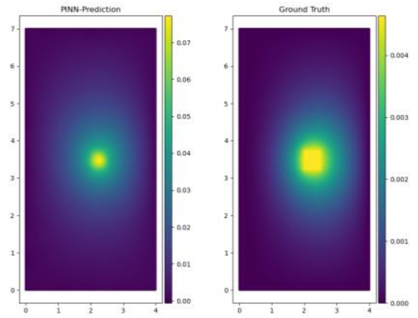
Use-Case a1

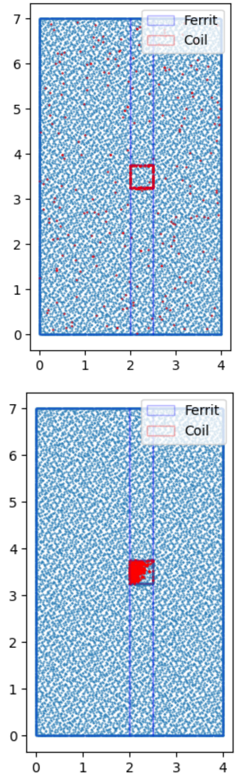






- early Trials without Interface Conditions:

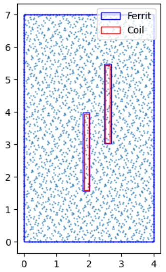


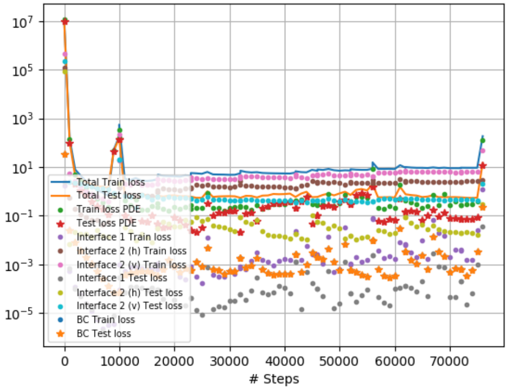


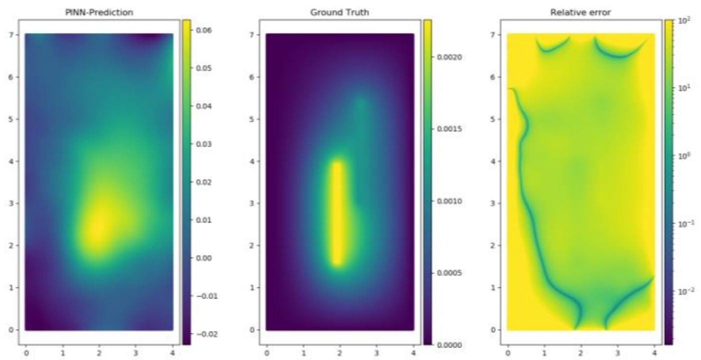
- Residual-based Adaptive Refinement

- To be continued…

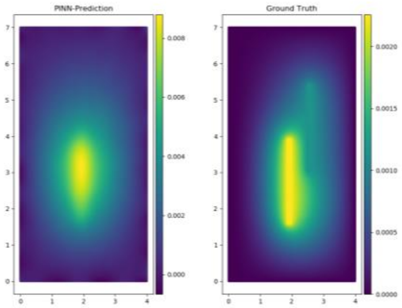
Intermediate Results
Use-Case b1

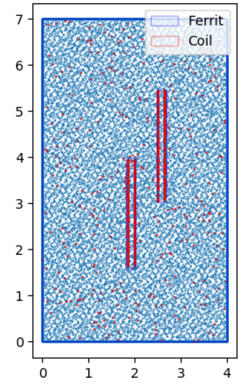


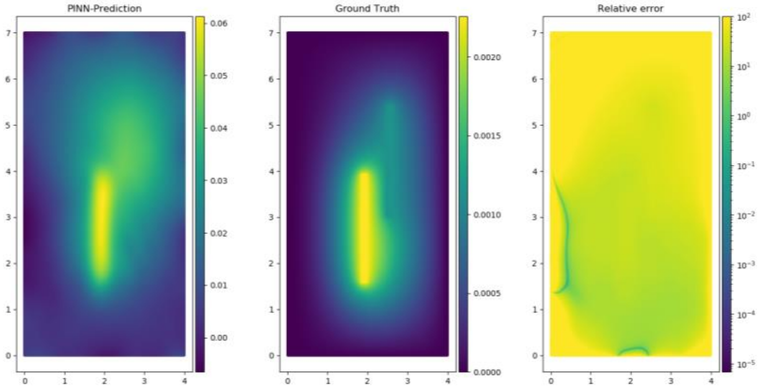




- early Trials without Interface Conditions:







- Residual-based Adaptive Refinement

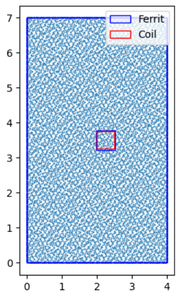
- To be continued…

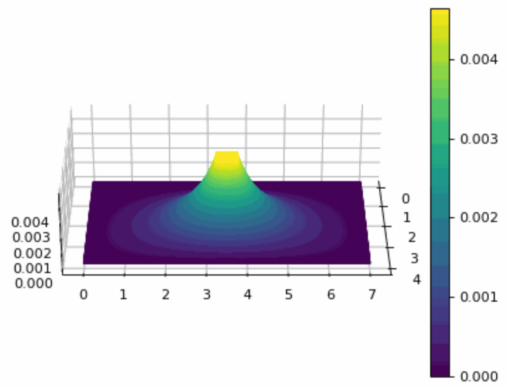
Intermediate Results

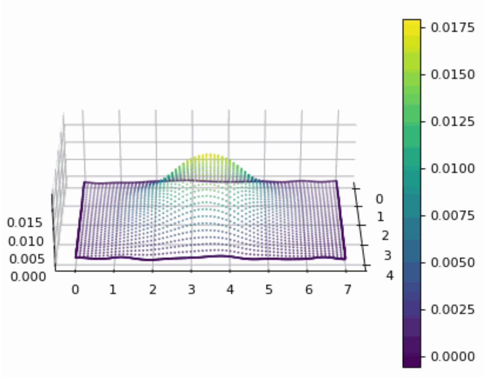
Effect of Interface Conditions

$$
\begin{gathered}
(A_{z2} - A_{z1})n_y = 0,\ (A_{z2} - A_{z1})n_x = 0 \text{ auf } \Sigma_{12} \\
\left(\frac{1}{\mu_1}\frac{\partial A_{z1}}{\partial x} - \frac{1}{\mu_2}\frac{\partial A_{z2}}{\partial x}\right)n_z = 0,\ \left(\frac{1}{\mu_1}\frac{\partial A_{z1}}{\partial y} - \frac{1}{\mu_2}\frac{\partial A_{z2}}{\partial y}\right)n_z = 0, \\
\left(\frac{1}{\mu_2}\frac{\partial A_{z2}}{\partial x} - \frac{1}{\mu_1}\frac{\partial A_{z1}}{\partial x}\right)n_x + \left(\frac{1}{\mu_2}\frac{\partial A_{z2}}{\partial y} - \frac{1}{\mu_1}\frac{\partial A_{z1}}{\partial y}\right)n_y = 0 \text{ auf } \Sigma_{12} \\
\sigma_1 n_z \overline{A}_{z1} - \sigma_2 n_z \overline{A}_{z2} = 0 \text{ auf } \Sigma_{12}.
\end{gathered}
$$

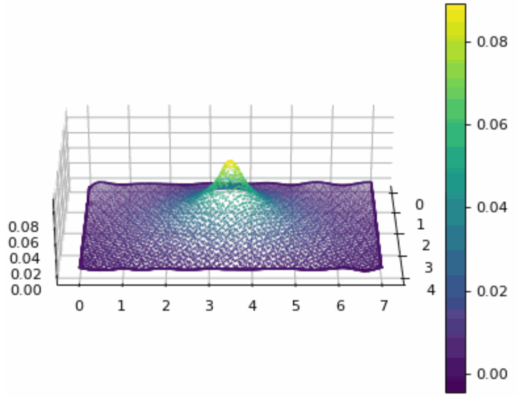
- Use-Case & Reference







- without Interface Conditions



- with Interface Conditions

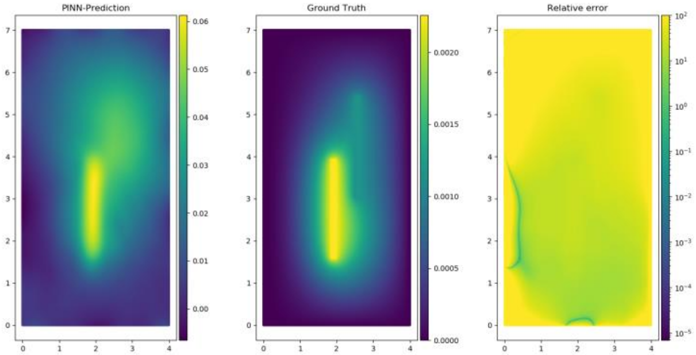
Outlook

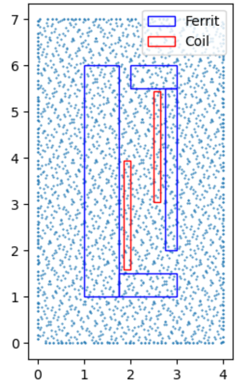
- Implementation & Evaluation of more advanced PINN-Specific Methods
  - Augmented Lagrangian Method for Loss-Balancing
  - Good Strategies for adaptive Sampling

- Tuning of Hyperparameters:
  - choice of optimizer (ADAM, L-BFGS,
  - activation functions, learning rate(-decay)
  - neural network size and structure
  - Parameter related to training schemes → Residual-based Refinement, Lagrange-Method
  - …there are many!

- Parametrization of Topologies → Generalization of PINN for arbitrary Topologies

- Further Read → Recent Survey on PINNs for Scientific Machine Learning [18]





References

Images:

- [1] Planar Inductor
- [2] Ferrit-Cores
- [3] Lithography

Publications & Software:

- [4] Raissi, Perdikaris, Karniadakis: Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations, 2017
- [5] DeepXDE: List of Publications
- [6] Nvidia Modulus: User Guide, Release v21.06
- [7] Choudhry: A Novel Framework for Physics based Machine Learning in Science & Engineering Domains, 2021
- [8] Angermeier, Stingl, Rosskopf: Topologieoptimierung induktiver Komponenten, Master-Thesis, 2019
- [9] Harlander: Numerische Berechnung von Induktivitäten in dreidimensionalen Verdrahtungsstrukturen, PhD, 2002
- [10] Biro, Preis: On the use of the magnetic vector potential in the finite element analysis of 3D eddy currents, 1989
- [11] Acevedo, Loaiza: A Fully-Discrete Finite Element Approximation for the Eddy Currents Problem, 2013
- [12] Beltran-Pulido et al.: Physics-informed neural networks for solving parametric magnetostatic problems, 2022
- [13] Hennigh et al.: NVIDIA SimNet™: an AI-accelerated multi-physics simulation framework, 2020
- [14] DeepXDE: A deep learning library for solving differential equations, 2021
- [15] FEMM: Finite Element Method Magnetics
- [16] Wang, Yu, Perdikaris: "When and Why PINNs fail to Train: A Neural Tangent Perspective", 2020
- [17] Lu, Pestourie et al.: Physics-Informed Neural Networks with hard constraints for inverse Design, 2021
- [18] Cuomo et al.: Scientific Machine Learning Through Physics–Informed Neural Networks: Where we are and What's Next, 2022

Fraunhofer IISB

Thank you for your Attention!

Fraunhofer IISB

Fraunhofer-Institut für Integrierte Systeme und Bauelemente-technologie IISB