

Master Thesis

Title: Element Identification in PGNAA spectra with Artificial Neural Networks

Speciality: Students of Mathematics, Physics or Related Fields

Motivation: Prompt Gamma Neutron Activation Analysis (PGNAA) is a measurement technique that enables material characterization in a non-destructive way. Applications include resource exploration, environmental analysis and improvement of recycling processes. Hereby neutrons are irradiated on a sample and induce the emission of gamma radiation. The evaluation of the resulting gamma spectra is a challenging task.



Subject: Within joint research and development projects we support our industrial partner AiNT in the development of PGNAA measurement facilities. For the practical application, it is necessary to evaluate the gamma spectra with high computational efficiency and accuracy. Hereby it is of great importance to identify peaks in the spectra and assign them to specific elements since these peaks carry the information about the material composition of the sample. The objective of this thesis is to develop a neural network that can identify the elements within the sample using its PGNAA spectrum.

Tasks: The work on the thesis will include following steps:

- Determination of a suitable network architecture.
- Implementation of the neural network using modern machine learning libraries.
- Training of the neural network with the provided data.
- Verification of the results through comparison with existing methods.

Contact: Prof. Dr. Martin Frank, Alexander Jesser Faculty of Mathematics Steinbuch Centre for Computing www.scc.kit.edu Email: martin.frank@kit.edu, alexander.jesser@kit.edu



Title: Diagonally preconditioning collocation methods

Speciality: Students from the Department of Mathematics

Motivation: Collocation methods are numerical techniques commonly used for solving differential equations and integral equations. These methods involve selecting a set of points, known as collocation nodes. By satisfying the equations at these chosen locations, collocation methods provide accurate high-order approximations for the solutions of a wide range of mathematical and physical problems.

When solving the collocation problem that is not represented by a triangular matrix, the process may take a long time to compute. In hope to speed up this process, we can obtain a solution by deploying preconditioned Richardson iterations where the preconditioning matrix is a diagonal one. Because of this, each iteration is not only straightforward to solve but also introduces parallelism across the collocation nodes.

Subject: The work on the thesis is concerned about finding a diagonal preconditioner that ensures convergence, and more importantly, it minimizes the number of iterations. It is not particularly clear how to define the diagonal matrix for different choices of Δt . In addition, the preconditioner will also dependent on the number of collocation nodes as well as on different distributions, i.e. the Radau IIA, Lobatto, Gaussian, etc.

- To set up and implement the collocation problem solving the test equation, the linear advection and the linear heat equation, carefully choosing numerical parameters in order to reach (for example 3) pre-chosen accuracies
- To come up and systematically test different diagonal preconditioners and the number of iterations the fixed point iterations need until convergence
- To compare the findings to already proposed preconditioners in the literature
- To examine the stability of the diagonal preconditioners after a fixed amount of iterations (for example 3 or 5, or until reaching a given tolerance)
- Contact: Gayatri Caklovic, PhD (SCC/SCM, Department of Mathematics) Prof. Martin Frank (SCC, Department of Mathematics) Email: gayatri.caklovic@kit.edu



Master Thesis

Title: Multi-label (N)TCP profile optimization on sparse imbalanced data for radiotherapy treatment planning

Courses: Mathematics, Computer Science, or similar

Motivation: Outcome predictions from normal tissue complication probability (NTCP) machine learning models can serve as quantities of interest in the optimization of a radiotherapy treatment plan. Typically, such models are trained on a single viewpoint, e.g. early complications within 6 months or long-term complications within 15 to 24 months after treatment, and thus the generated plan only minimizes the risk of complications during that time period. A potential advancement could be the transition to a multi-label case, where NTCP is modeled for a sequence of observation time points and then optimized, effectively representing the optimization of an NTCP profile.



Subject: The work on this thesis is an extension of ongoing research regarding the incorporation of ML-based objectives and constraints into optimization problems, and will include the topics of sparse imbalanced data handling, multi-label classification, and vector optimization (Pareto optimization). The student is expected to develop application-oriented strategies for optimizing and interpreting NTCP profiles and implement these within the framework of our Python package *pyanno4rt*.

Tasks: The following steps could be part of your work:

- Familiarize yourself with the field of vector-valued NTCP modeling and optimization in the context of radiotherapy treatment planning
- Design and implement machine learning NTCP profile prediction models in view of sparse imbalanced head-and-neck patient data
- Develop optimization strategies for NTCP profiles, e.g. based on mean or maximum NTCP, or more sophisticated approaches
- Analyze your concepts within the *pyanno4rt* environment and extend the modeling and visualization suite of the package

Contact: Tim Ortkamp, MSc (SCC/SCM, DKFZ) Prof. Martin Frank (SCC, Department of Mathematics) Email: tim.ortkamp@kit.edu



Bachelor's or Master's Thesis

Title: Monte Carlo integration with artificial neural networks

Speciality: Students from the Department of Mathematics

Motivation: A task of modern particle physics is to verify (or disprove) our understanding of the laws of nature. To that end, experiments in particle accellerators, e.g. at the Large Hadron Collider at CERN, are conducted. Theoretical predictions are (differential) cross sections, i.e. certain probability distributions, and there is a need for (numerical) integration to obtain testable probabilities. Deterministic quadrature algorithms suffer from the "curse of dimensionality" (number of integrand evaluations grows exponentially with the dimension). This is mitigated by the use of Monte Carlo integration. Modern approaches utilize artificial neural networks (ANNs) for a more efficient sample generation.



A tunnel section of the LHC Source: <u>https://commons.wikimedia.org/wiki/File.CERN_LHC_Tunnel1.jpg</u> Author: <u>Julian Herzog</u>



An illustration of Monte Carlo integration Source: <u>https://commons.wikimedia.org/wiki/File:MonteCarloIntegrationCircle.svg</u> Author: <u>Mysid Yaderi</u>

Tasks: Depending on the targeted degree, the work on the thesis will include some of the following steps:

- Document the ideas of Monte Carlo integration (close gaps in the proofs).
- Conduct numerical experiments with classical Monte Carlo integrator in Python using a modern framework on a toy example.
- Implement and document an ANN based Monte Carlo integrator in Python. Test it on a toy example.
- Apply an ANN based MC integrator to a more realistic integrand from particle physics.

Contact: Dr. Leonid Chaichenets (SCC/SCM, CoMMAS) Prof. Martin Frank (SCC, Department of Mathematics) Email: leonid.chaichenets@kit.edu Phone: 0721 608-44867



Title: Symmetrized Polynomial Propagators for Hamiltonian Systems

Field of study: Mathematics

Description: Dynamics of Hamiltonian systems is simulated using numerical time integration methods specific for this type of ordinary differential equations. Typical applications are molecular dynamics and celestial mechanics simulations. These applications require simultaneously several properties from time integrators: sufficient order of accuracy, long-term stability, low computational complexity and high scalability and short time to solution. The small and fixed time steps necessary to fulfil some of the requirements limit the success of currently used integrators. Expanding the time evolution operator in series of Chebyshev or Newton polynomials provides promising schemes, allowing larger time steps and parallelization. On one hand, these schemes are insufficiently studied in the case of classical dynamics, and on the other hand, they do not preserve the intrinsic time reversal symmetry of the solution required for long-term stability.

Subject: For a small test system (with three particles), you will investigate one-step implicit time-symmetric schemes based on the original polynomial schemes focusing on the local truncation error (LTE), the global error and their relation to the time step size and polynomial order.

Tasks:

- Investigate how the LTE of the two original schemes and of their symmetrized versions behave with variation of expansion order and time step.
- Show analytically and numerically that the Chebyshev propagator is time-symmetric in the limit of infinite expansion order and compare to the Newtonian propagator.
- Measure the global energy error (energy drift) for the original and the symmetrized schemes again with increasing expansion order and time step.
- Decompose the LTE of the symmetrized scheme into components accounting for the asymmetry and for the series truncation error, and study their behavior in the limit of very short time step size using sufficiently high decimal precision.
- Suggest a modification of the symmetrized scheme allowing adaptive time step size using a suitable criterion for the LTE.

Contact: Dr. Ivan Kondov, Prof. Dr. Martin Frank Email: ivan.kondov@kit.edu Phone: 0721 608-28644



Bachelor or Master Thesis

Title: Comparative study of rank-adaptive DLR integrators

Specialty: Students from the Department of Mathematics

Motivation: Problems with a high dimensional phase space pose several challenges for scientific computing and simulation. For instance, computational and memory resources required for resolving solutions increases exponentially with dimensions. In the recent years dynamical low-rank approximation (DLRA) has gained popularity as a method to address these challenges for relevant problems, like radiation transport problems. DLRA is a numerical method that approximates the solution on a low-rank manifold by projecting and evolving the dynamics on the tangent space to the low-rank manifold. Since the rank of the solution is not known beforehand several integrators, that adaptively select the rank of the approximation, have been developed. However, it is not entirely clear how they perform in comparison to one another.



 $F(Y(\tilde{t}))$ $Y(\tilde{t})$ Y(t) $T_{Y(\tilde{t})}M_{r}$ M_{r}

Linesource testcase with BUG integrator

Projection onto tangent space to low-rank manifold

Subject: The work on the thesis is connected with the development of robust DLRA integrators for radiation transport problems. For this type of integro-differential equation, since it's a hyperbolic problem, the solution develops shocks and discontinuities. Furthermore, deriving evolution equations for the low-rank approximation as well as designing experiments on the basis of which all the integrators are to be compared; these aspects make it a challenging task.

- Do a literature review of dynamical low-rank approximation as well as radiation transport equations.
- To derive and implement various fixed-rank and rank-adaptive integrators from literature for the radiation transport equation and run 1D and 2D test cases for each integrator.
- Select the metrics and design experiments to compare the various rankadaptive integrators available in the literature.
- Possibly develop a new rank-adaptive integrator that addresses the possible short comings faced by the other integrators.
- Contact: Chinmay Patwardhan, M.Sc. (CSMM, CRC 1173) Prof. Martin Frank (SCC, Department of Mathematics) Email: chinmay.patwardhan@kit.edu



Title: Uncertainty quantification on laminar flow

Speciality: Students from the Department of Mathematics

Motivation: Laminar flows are fundamental in engineering and science. Uncertainties in parameters, initial and boundary conditions can significantly affect computational fluid dynamics (CFD) predictions. Developing reliable uncertainty quantification methods is essential. These methods enhance CFD simulation reliability and improve safety in practical applications, optimizing designs, and providing deeper insights into laminar flow behavior. The study's motivation is to advance CFD by developing efficient, accurate uncertainty quantification methods, aiding decision-making across various applications.



Velocity contour within Monte Carlo method

Subject: In this project, we aim to utilize the lattice Boltzmann Method to simulate laminar flows, with a specific focus on scenarios like flow past a circular cylinder. Uncertain parameters in the inlet boundary conditions and initial conditions present a challenge as deterministic methods fail to provide reliable results. Therefore, our objective is to incorporate an uncertainty quantification framework into our model. Several key questions remain to be addressed, including the identification of uncertain parameters, the determination of the probability distributions governing these parameters, and the effect of uncertainty on the quantity of interest.

Tasks: The work on the thesis will include following steps:

- To setup an OpenLB model in a 2D or 3D geometry.
- Implement the Monte Carlo method or the stochastic collocation method
- Quantify the uncertainty

Contact: Mingliang Zhong Prof. Martin Frank (SCC, Department of Mathematics) Email: mingliang.zhong@kit.edu



Bachelor/Master Thesis

Title: Uncertainty Assessment of Stochastic Optimization Algorithms for Battery Parameter Estimation

Speciality: Mathematics or Computer Science

Motivation: Monitoring and controlling the battery play a critical role in the development and performance of electric cars, which drew attention to intelligent battery management systems (BMS). Enhancement of the battery management systems using AI methods requires an adequate model of the battery and the most promising approach is equivalent circuit model (ECM). The equivalent circuit is the electrical counterpart of an electrochemical system. A non-convex inverse problem should be solved to find the equivalent circuit of a battery using measurement data and optimization algorithms are used to approximate the analytical solution.



Non-convex search space(egg holder function)

Subject: The work on the thesis is to analyse the uncertainty caused by the randomness in stochastic optimization algorithms that are favored for non-convex search spaces. The implications of this uncertainty is not trivial and different simulation runs may return different local minima with almost equally importance with respect to the cost function. The choice of the solution and the stability of the algorithm are strongly connected to the quality of equivalent circuit.

Tasks: The work on the thesis will include following steps:

- To examine the optimization problem mathematically and compare most common stochastic optimization methods according to search space analysis for a simple circuit (Bachelor/Master)
- To choose one or two promising methods and tune the stochastic parameters on a more complex circuit (Bachelor/Master)
- To carry out multiple runs with tuned parameters to get statistics of returned solutions and report if the observed variability has implications to achieve convergence to global minimum(Master)
- To do sensitivity analysis of the solutions to stochastic parameters(Master)

Contact: Hasan Hüseyin Kalkan(SCC/SCM) Prof. Martin Frank (SCC, Department of Mathematics) Email: hasan.kalkan@kit.edu



Title: Modeling the collective behavior of flexible fibers in Isotropic Forced Turbulence

Motivation: Particle-laden turbulent flows are multiphase systems in which a carrier fluid interacts with a dispersed phase composed of solid objects such as spheres or fibers. The collective behavior of small, flexible fibers can be highly intricate within turbulent flows. Previous studies on fiber dynamics have predominantly concentrated on rigid fibers or rods. Due to their elastic properties, the flexible fibers exhibit a unique behavior within particle-laden turbulent flows. Unlike rigid fibers or rods, the flexibility of these particles allows them to deform and stretch in response to the turbulent forces acting upon them. This flexibility enables them to absorb and store energy within the turbulent flow field, leading to intricate interactions between the fibers and the surrounding carrier fluid.

Subject: In this project, we are interested in using Euler–Bernoulli beam theory to develop a library in OpenFOAM for modeling flexible fiber particles. This library will be used later to study particle preferential concentration and the effect of these particles on turbulent flow characteristics.

Tasks: The thesis work will involve the following steps:

- 1- Develop the Lagrangian library of OpenFOAM to model rod-like particles.
- 2- Introduce flexibility using Euler-Bernoulli beam theory.
- 3- Integrate the developed code with an existing homogeneous isotropic turbulence solver.

Software skills: OpenFOAM; C++

Contact: Prof. Dr. Martin Frank, Hesam Tofighian Faculty of Mathematics Steinbuch Centre for Computing www.scc.kit.edu Email: martin.frank@kit.edu, hesam.tofighian@kit.edu



Master Thesis

Title: Uncertainty in the calculation of ozone due to use of different horizontal resolutions in a Earth System Model

Courses: Mathematics

Motivation: Uncertainties in Earth System Models are various. One of a larger uncertainty comes from the calculation of the chemistry, which is influenced by many factors (e.g. uncertainties in the laboratory measurements of the rate constants of the chemical reactions). One of these factors is the inaccurate calculation of the temperature due to a coarse resolution in the Earth System Model. In the case of a coarse resolution, the grid boxes are very large. Thus, the actually much finer temperature distribution of temperature is averaged away within one grid box and you get a linear mean. However, the calculation of the chemistry is done by the Arrhenius equation (or similar equations), which is not linear and in which the temperature is included exponentially. This causes deviations in the calculation of the chemical reactions. How large these deviations are, is not known and therefore task of this master thesis.

Tasks: The focus of the master thesis is the chemical substance ozone. The question is which influence have the different horizontal resolutions to the calculation of ozone with regard to the temperature distributions? Is it possible to calculate a correction for this calculation (resolution dependent)?

The work on the thesis will include following steps:

- Performing of about three one-year ICON simulations with comprehensive chemistry, each simulation with a different horizontal resolution (e.g., R02B05 (80 km), R02B07 (20km), and R02B09 (5km)).
- Evaluation of ozone and statistical determination of uncertainty with regard to the temperature distribution.
- Definition of a correction formula, which can be built into ICON.
- Performing of ICON simulations with the same coarse resolution and setup as before, but with integrated correction formula. Comparison with the original simulations.

Contact: Dr. Ole Kirner (SCC-SCM, SimDataLab Earth System Science) Prof. Dr. Martin Frank (SCC, Department of Mathematics) Email: ole.kirner@kit.edu Phone: 0721 608-25698



Title: Photonic materials with properties on demand designed with AI technology

Specialty: Students from the Department of Physics

Motivation: The ability to design photonic materials with predefined properties to mold the flow of light is an intellectual challenge and of utmost importance in many impact areas such as sensing, imaging, or emerging quantum technologies. Using artificial neural networks will accelerate the design process to a considerable extent.

Subject: This project uses artificial neural networks as surrogate models in an inverse design problem of finding nano-structured materials with optical properties on demand. In order to design the objects, the optical properties of the scatterers are represented parametrically. These parameters constitute the base for an algebraic description of the light-matter-interaction. In this description, incident and scattered fields are expanded into basis functions, and matrix multiplication connects both. The connecting matrix is called the T-matrix, and its size depends on the truncation order. The T-matrix is a frequency-dependent quantity, and it contains all the information on how an object scatters an incident field. For a given object, the T-matrix can be computed from the geometry and the material composition of an object by solving the full-wave 3D Maxwell equations.

Current state of the project: Large amounts of data from 3D simulations of Maxwell's equations (COMSOL, JCM wave), were generated for axiosymmetric scatterers. The resulting T-matrices were used for neural network training. Given the object, the network is able to reconstruct the corresponding T-matrix, based on object's geometry and material composition.

Tasks: The actual challenge is to solve the inverse design problem. This requires identifying a physically existing object from some given materials that offers a predefined T-matrix. Tailored search algorithms and code addressing the efficient inversion are to be developed. Once we know the most optimal T-matrix for a specific application, our second challenge concerns designing an object that offers the desired T-matrix. The design, ideally, should be constrained by existing materials and geometries accessible by available fabrication technologies to realize these scatterers.

Contact: Maria Paszkiewicz (SCC-SCM, FTP) Prof. Dr. Carsten Rockstuhl (FTP, INT) E-mail: maria.paszkiewicz@kit.edu



Title: Optimization of 2D waveguide trajectories using approximation methods for calculation of transmission

Specialty: Students from the Department of Physics

Motivation: Photonic wire bonds are freeform dielectric waveguides that connect optical chips made of different materials in fully integrated photonic devices at a microscopic scale. Utilization of approximate methods for the fast calculation of transmission has practical applications for the accelerated design of the wire bonds. In this case a numerical solution to Maxwell's equations is no longer feasible due to the long computation times.



Figure 1. Magnitude of the instantaneous electric field in a dielectric waveguide upon considering the fundamental mode of a straight waveguide as the illumination from the left end. Upper figure shows the optimal shape of the trajectory and lower figure shows the trajectory causing high transmission loss.

Subject: We have developed methods like fundamental mode approximation and multimode approximation which are suitable merit functions for trajectory optimization in which the transmission should be maximized. Using the approximation, we want to find optimal trajectory connecting input and output port, which avoids also all obstacles in between.

Tasks: The work on the thesis will include following steps:

- Exploration of different description of the paths to be optimized, e.g. series of Cartesian coordinates, series of arcs, Bézier curves, B-splines etc.
- Comparison of different optimization procedures w.r.t. number of parameters to be optimized, convergence, precalculation times, optimization times.
- Finding optimal trajectories for several test cases.

Contact: Maria Paszkiewicz (SCC-SCM, FTP) Prof. Dr. Carsten Rockstuhl (FTP, INT) E-mail: maria.paszkiewicz@kit.edu



Title: Use of scattering from an inhomogeneity inside a dielectric waveguide to overcome reflections in waveguide bends

Specialty: Students from the Department of Physics, KSOP

Motivation: Waveguides are used for transmission of electromagnetic power in photonic devices. Physical constraints in such structures sometimes require bending of the waveguides along their path, which worsen their efficiency due to reflection losses. Over the past few decades there have been several ways proposed in the literature to overcome these losses. The understanding of scattering behaviour of the electromagnetic field can be used to overcome losses in waveguides induced by bending.

Subject: The idea of this master's thesis is to modify the scattering properties of the bend by placing a single polarizable element (a scatterer) inside the bending region of a waveguide.

The interaction of guided modes with scatterers placed inside dielectric waveguides can be described analytically for slab waveguides and/or for fiber waveguides. The interplay of propagating waves and the obstacle is described in terms of forward and backward scattering coefficients and mode profiles of TE (transverse electric) and TM (transverse magnetic) fields.

One interesting practical problem in this respect is the ability to predict the effects of impurities inside the dielectric waveguides. Impurities can be caused by material defects, by captured air bubbles inside the material, or by the junction of two dielectric waveguides, etc. The analytical nature of the solution furnishes interesting results, such as mode-coupling effects, resonance phenomena, and rotation of polarization plane.

- Study of the topic and reproducing the results available in the literature.
- The analytical solution shall be compared with full-wave simulation in COMSOL or CST Microwave Studio.
- Finally, the effect of scatterer placed in different kind of bends (e.g. smooth bend with varying bend angle) should be analyzed.
- Contact: Maria Paszkiewicz (SCC-SCM, FTP) Prof. Dr. Carsten Rockstuhl (FTP, INT) E-mail: maria.paszkiewicz@kit.edu



Title: Soot particle image segmentation and analysis of particle agglomerates using deep learning models

Speciality: Students from the Department of Mathematics

Motivation: Primary particle size analysis is used to analyse particle size and particle size distribution. There is a wide range of applications where, in the project, the particle size analysis of soot agglomerates is performed. Transmission electron microscopy (TEM) images of soot agglomerates are produced. With the image segmentation and analysis, along with state-of-the-art imaging particle size methods like Hough transformation, ultimate erosion, and watershed transformation, deep learning models are used for the analysis.



Subject: The exhaust gas emissions of internal combustion engines are observed more during the cold start condition of passenger vehicles. The number of particles emitted during the cold start increases significantly, and under the influence of adsorption and condensation effects, a powdery soot layer is observed and formed on the walls of the exhaust gas flow. The soot particles are collected, and transmission electron microscopy (TEM) images of soot agglomerates at different engine cold start temperatures are captured. The goal of the thesis is to perform effective image segmentation and analysis of the primary particles of soot agglomerates.

- Pre-processing of images with soot agglomerates.
- Data preparation for segmentation task with recognition of the circular soot particles in the agglomerates.
- Image processing, implementation of the neural networks, training, validation and analyzing the particle size distributions.
- Contact: Manoj Mangipudi, MSc (SCC/SCM) Prof. Henning Bockhorn (KIT, Engler-Bunte Institute) Email: venkata.mangipudi@kit.edu



Master Thesis or Bachelor Thesis

Title: Data Preprocessing and AI Training

Speciality: Students from the Department of Mathematics

Motivation: During a previous PhD thesis, engine parameters (over 200 quantities) have been thoroughly recorded experimentally during cold start from different initial temperatures of an IC engine. The data have been collected from different devices with different frequency rates and need first to be preprocessed and collected together. Then, they will be used for training of a feed-forward neural network (NN). There are more than 100 separate measurements with different initial temperatures and under different driving conditions. The Figure below shows some of the engine parameters recorded during one of the cold start measurements at 5 Hz frequency.



Subject: The cold start of an engine is characterized by increased hazardous emissions and the release of fine particles. Therefore, a special focus is put exactly on the first 10 minutes of the driving cycle and only these data will be processed in the thesis. With a frequency range of 5 Hz this results in about 3000 time-steps. As such measurements are quite expensive, the idea of the thesis is to predict the hazardous output during new (not measured previously) cold start driving conditions. For this purpose, a relatively classical type of neural network (feed-forward NN) will be utilized.

Tasks: The work on the thesis consists of two parts – data preparation and AI training with these data. The first part requires adjustment of the time-series from the different measuring devices. Here, in order to match the data from the devices in time, first a correlation coefficient needs to be computed for different time-lags. After the data are matched according to the time-lag and the sampling frequencies, they will be stored in a new file (one file per measured case). With these pre-processed data, a neural network (NN) will be trained and tested. This neural network will be created, trained and validated using python/pytorch utilities.

Contact: Jordan Denev, PhD (KIT, SCC/SCM, SimDataLab EEM) Prof. Henning Bockhorn (KIT, Engler-Bunte Institute) Email: jordan.denev@kit.edu Phone: 0721 608-25771