

An AI-Based Approach to Optimize Stress in Shrink Fits

V. Dausch, J. Kröger  and M. Kreimeyer

University of Stuttgart, Germany

 jan.kroeger@iktd.uni-stuttgart.de

Abstract

The present analytical design of shrink fits typically results in an uneven stress condition that can lead to failure in a variety of manners. With increasing loads and the use of brittle materials, the optimization of the stresses in the shrink fit hence becomes increasingly necessary. Currently existing approaches do not solve the problem satisfactorily or increase the manufacturing and design effort. This paper therefore considers the implementation of an AI-based stress optimization using reinforcement learning, which performs stress optimization by geometrically contouring the interstice.

Keywords: artificial intelligence (AI), engineering design, numerical methods, optimisation, structural analysis

1. Introduction

Shaft-hub connections (SHCs) are an important machine element (Kollmann, 1984) in many drive applications. Cylindrical shrink fits are frequently used due to their simple production, the resulting low manufacturing costs and their resistance to shocks and excessive stresses (Leidich, 1983; Kollmann, 1984). The transmission of power between hub and shaft takes place via a frictional connection which usually is caused by elastic deformation (Leidich, 1983; Kollmann, 1984; Kröger et al., 2020). The demands on SHCs to optimize performance, durability and weight are continually rising - if this is achieved by significantly increasing the extent to which the properties of the materials are utilized, this means higher stresses on the components (Kröger et al., 2020). High pressures within the joint and/or brittle materials can then cause failure of the joint. The way to avoid this is to employ complex optimization techniques; these require in-depth expert knowledge and time-consuming FEM simulations. By using AI, an automated design of the joint leads to a time-optimized and less complex optimization for the user. The purpose of this article is to summarize the possibilities, the approach taken so far, and the current research results carried out at the research center with regard to AI-assisted design of shrink fits.

2. Problem statement and objective

The design of cylindrical shrink fits as specified in (DIN, 2017) is based on a shrink fit that is radially and axially symmetrical, with a shaft and hub of the same length. This leads to the assumption of a plane stress condition (psc), i.e. a constant and even distribution of radial and tangential stresses across the whole joint, for which the constant pressure profile at the interstice of the connection is shown in (Figure 1c). With the simplified assumption of a plane stress condition, the design can be performed using analytical calculation equations. However, this simplified example of a shaft and hub of identical length does usually not correspond to practical reality, which can lead to an overestimated safety coefficient. This is because there is inevitably a jump in stiffness at the ends of the hub. This jump cannot be solved

analytically and/or numerically in a precise manner and results in an uneven pressure profile within the joint (cf. Figure 1d). Peaks or general variations in pressure within the joint can have various consequences for the joint. For brittle materials such as cast materials or heavily carburized case-hardened steels, high stress peaks can cause early spontaneous failure (Blacha, 2009; Glöggler, 2003; Hartmann, 1999; Rentzsch et al., 2002). If due to stress variations the pressure within the joint drops locally, the slip amplitude under operating loads inevitably increases. In the long term this can lead to failure due to fretting corrosion (Leidich, 1983; Gropp et al., 2012; Vidner, 2016). In the event of an extreme drop in pressure within the joint, the joint may even slip. Previous research work, such as (Blacha, 2009; Krautter et al., 2015; Kröger et al., 2018), has already developed a variety of solutions for the optimization of shrink fits. However, these approaches are not yet suitable for general application due to their high complexity and lack of general calculation approaches; they often represent only approximate iterative solutions, and their application is time-consuming. Therefore, this article presents an option for an AI-supported optimization process that especially increases the efficiency of R&D-processes. The main research question of this article is how the optimization process of the pressure within the joint - or stress optimization in general - of shrink fits can be supported using AI. This main research question can be subdivided into more detailed sub-research questions:

- Could AI be used to optimize the material utilization and the stress conditions of shrink fits?
- Which algorithm is suitable for this use case?
- Is it possible to reduce the time required for optimization with the help of AI?
- Does an AI supported optimization provide the desired quality of results?

3. Shrink fits

3.1. Stress condition on cylindrical shrink fits

The design of cylindrical shrink fits is specified in (DIN, 2017). They can be designed either linear-elastic or elastic-plastic. Due to a lack of experimental validation, linear-elastic shrink fits are generally preferred in industrial practice (Kröger et al., 2020). With this type of connection, the transmission of forces and torques is based on a frictional connection created by a geometric interference between the outside diameter of the shaft and the inside diameter of the hub (Kollmann, 1984). The elastic deformations that occur when this interference is overcome result in a triaxial stress condition (Leidich, 1983). In that stress condition there is no proportionality between the radial stress σ_r and the tangential stress σ_t over the interference U within the overall shrink fit joint. The radial stress σ_r corresponds to the pressure within the joint p in the area of the interstice and decreases towards the outside. The elastic expansion of the hub also results in a tangential stress σ_t (cf. Figure 1a). In the axial direction, the joint is mainly stressed by shear stresses τ_a induced by friction (Figure 1b). (Leidich, 1983)

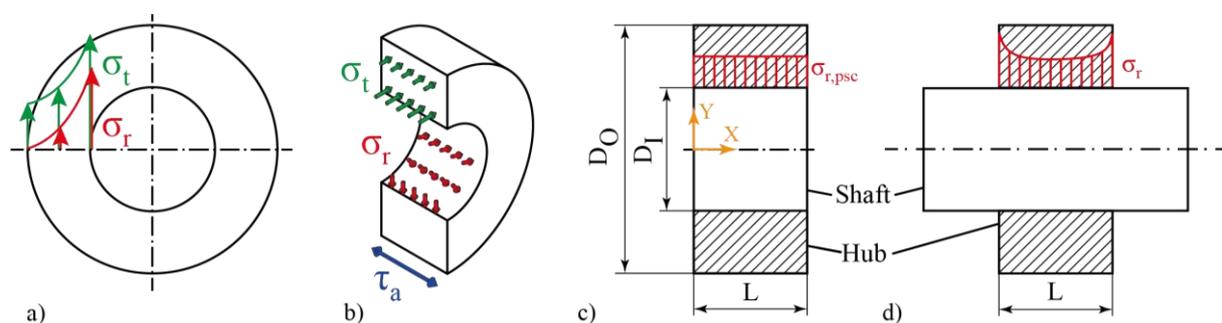


Figure 1. Principal stresses on shrink fits in transversal a) and longitudinal section b); radial stress in the interstice of shrink fits within plane stress conditions c) and triaxial stress state d)

In the following equations, stresses of the plane stress condition are generally referred to as σ_{psc} . The pressure within the joint according to (DIN, 2017) is also referred to as p_{psc} . For the radial (Equation 1) and tangential stress (Equation 2), the general rule according to (Kollmann, 1984) is:

$$\sigma_r(Y) = p_{psc} \cdot \left(\frac{D_I^2}{D_O^2 - D_I^2} \right) \cdot \left[\left(\frac{D_O}{2 \cdot Y} \right)^2 - 1 \right] \quad (1)$$

$$\sigma_t(Y) = p_{psc} \cdot \left(\frac{D_I^2}{D_O^2 - D_I^2} \right) \cdot \left[\left(\frac{D_O}{2 \cdot Y} \right)^2 + 1 \right] \quad (2)$$

At the interstice ($Y = D_I/2$), the radial (Equation 3) or tangential (Equation 4) stress is thus given by:

$$\sigma_r = -p_{psc} \quad (3)$$

$$\sigma_t = p_{psc} \cdot \frac{D_O^2 + D_I^2}{D_O^2 - D_I^2} \quad (4)$$

Due to the different lengths of shaft and hub in industrial practice, there is always a jump in stiffness at the ends of the hub, so that excessive stresses occur within that area. Because the analytical design is based on the simplified assumption of a plane stress state, it disregards the stresses that occur in reality (see Figure 1c). This becomes particularly critical when this method is used to design highly stressed shrink fits or connections using brittle hub materials. If the tangential stress σ_t reaches the separation strength of the hub material it will suffer spontaneous brittle failure. (Figure 1d) shows an example of the pressure profile within the joint of a practical shrink fit normalized to the pressure within the joint resulting from the plane stress condition. This means that basically all parameters (e.g. D_I/D_O , L/D_I) that influence the stiffness of the shrink fit can also influence the stress increase. The results in (Figure 2) based on research completed in (Kröger et al., 2018) illustrates the effects of different geometric dimensions on the stress increase.

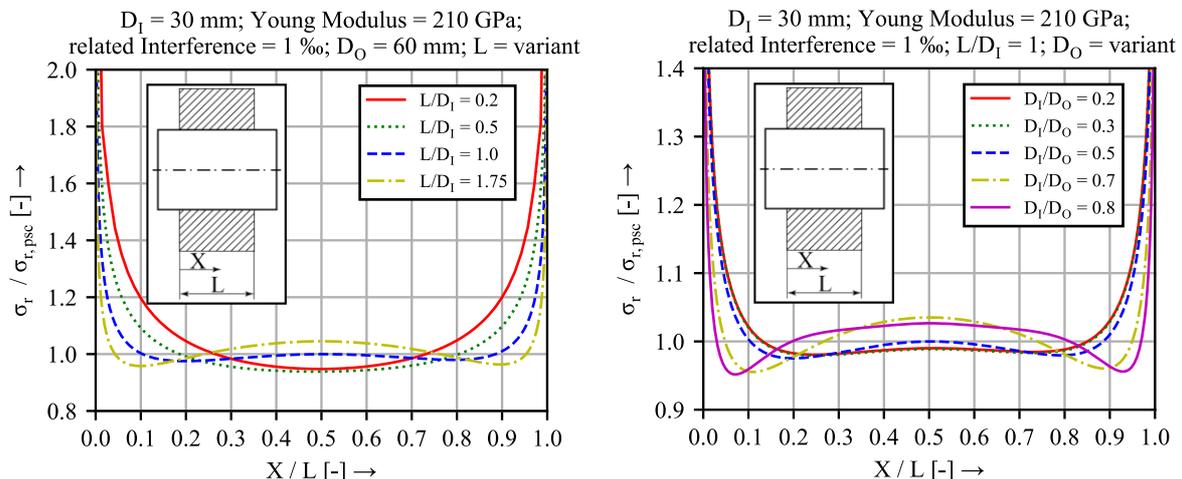


Figure 2. Influence of the length-to-diameter L/D_1 ratio (left) and the diameter ratio of the hub D_1/D_0 (right) on the radial stresses in the interstice based on (Kröger et al., 2018)

3.2. Simple methods to reduce stress peaks in the area of the hub ends

In literature (Leidich, 1983; Kollmann, 1984), there are many approaches aimed at enabling a reduction of the excessive stresses in the areas of the ends of the hub. Examples include circumferential grooves in the hub faces (cf. Figure 3d) and chamfers or radii at the inner hub edges. However, investigations (Leidich, 1983; Kollmann, 1984; Kröger et al., 2018) have shown that many of these approaches fail to improve the stress condition or suffer from other undesirable disadvantages. Using the examples of the circumferential groove on the hub face and the radius at the edge of the hub, (Figure 3) (left) illustrates the problems arising from simple geometric adjustments for stress optimization. Although the circumferential face groove effectively reduces the value of the stress peaks, it also reduces the pressure within the joint so that the inevitable increase in slip amplitude creates conditions favorable to the occurrence of fretting fatigue. In contrast, the use of a radius at the hub edge only causes an axial displacement of the location of the jump in stiffness and hence of the location of the increased stress. Other strategies, such for instance as those proposed by Kollmann in (Kollmann, 1984) or Leidich in (Leidich, 1983), create similar problems (Figure 3a). They require additional axial design space (Figure 3a, b) or restrict the common functional areas of the hubs with respect to their design (Figure 3b, c).

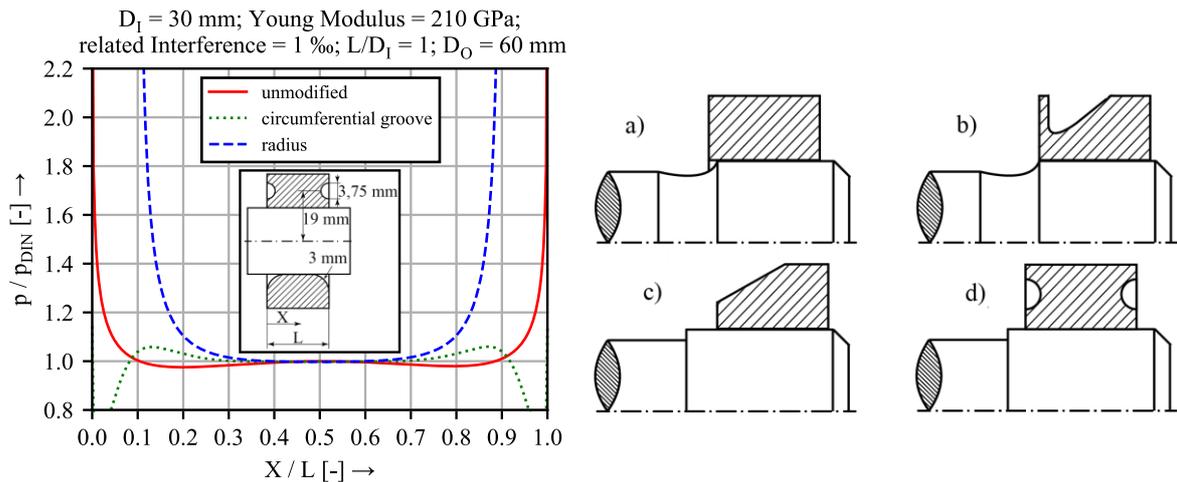


Figure 3. Comparison of the radial stress in the interstice of shrink fits with and without simple geometric modifications (left) and examples of simple geometric modifications (right, a to d)

3.3. Homogenization of the pressure within the joint through contoured shaft

Compared with the simple methods presented in section 3.2, complex contouring of the outer surface of the shaft and/or the inner surface of the hub can be used to prevent undesirable variations in stress (see Section 3.1). This is in fact the only way to achieve stress and pressure profiles within the joint that counteract failures due on the one hand to overstressing and on the other hand to frictional fatigue loading. In addition to the structural and tribo-mechanical advantages of this process, it also demands no additional space and creates no difficulties regarding integration into conventional CNC production.

Blacha mentions in (Blacha, 2009) that the same reduction in stress peaks (cf. section 3.1) for a shrink fit can be achieved by use of the pressure within the joint p or the tangential stress σ_t as input variables for the optimization process. In addition to (Figure 2), the necessity for a specific adjustment is also shown in (Figure 4) by means of comparison of the actual stress profile of an asymmetric shrink fit against the stress profile of the plane stress condition which was described in Section 3.1. The profile shows that complex optimization is necessary, in particular for asymmetric shrink fits (e.g. pulleys or gears with a collar), since the effects that occur at the two ends of the hub can be in the opposite sense to each other.

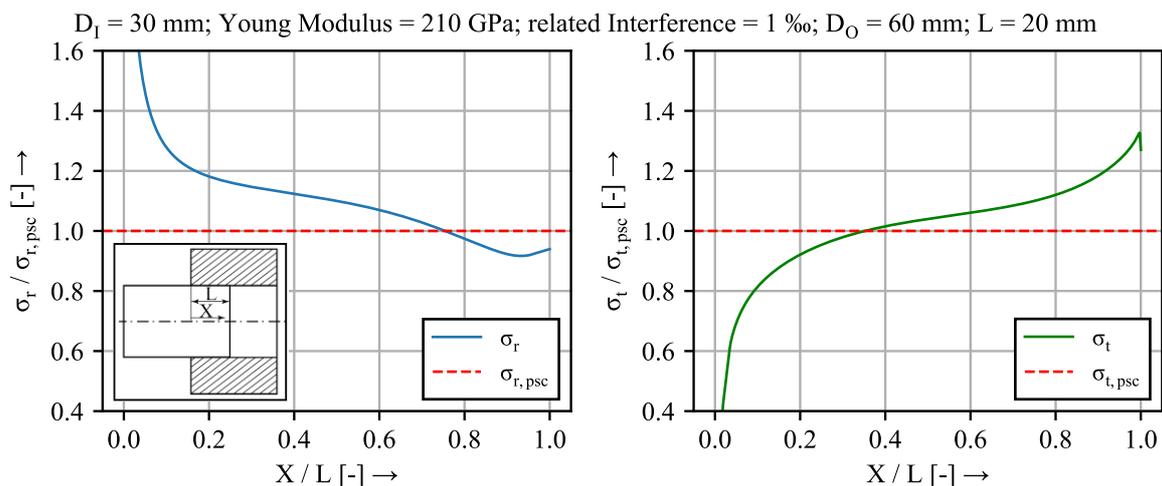


Figure 4. Excessive radial (left) and first principal stress (right) in asymmetrical shrink fits

At the research center, several methods (Blacha, 2009; Krautter et al., 2015) have already been developed for achieving a specific pressure within the joint. Basically, the method used by Blacha is based on a separate numerical simulation of shaft and hub, whereby the actual contact surfaces are subjected to the desired pressure profiles within the joint, which for example can be constant. By summing the amounts of the resulting component deformations, the desired contouring is obtained and

can be applied to the shaft or hub. For the ideal case of a frictionless interstice, this results in a constant profile of the pressure within the joint. To take account of the frictional forces within the interstice, Blacha includes these as additional linear boundary conditions in the individual simulations described above. In (Krautter et al., 2015), Krautter was able to show that the friction force profile in the interstice is non-linear. Krautter then extended the procedure with a thermo-mechanical simulation in order to determine the frictional force profile more precisely. What the existing methods have in common is that the procedures are very complex and time-consuming and also require expert knowledge. In addition, the optimization focuses solely on the pressure within the joint, so that it fails to offer precision in addressing what is the critical criterion for failure, i.e. the optimization of the main stress.

3.4. Concretization of the problem and requirements for an AI

Within certain limits, ductile hub materials can reduce the excess stresses (cf. Section 3.1) by local plasticizing, so that they do not have a negative effect in further operation. Brittle materials (e.g. cast materials, technical ceramics, composites or hardened components) can reduce these stress peaks to only a limited extent, hence their use can lead to failure. In addition, the trend toward higher utilization of material properties in terms of higher power densities and lightweight design causes reductions in the existing safety margins previously contributed by the ductility of hub materials; thus failures due to excessive stress can occur here also. Simple approaches to reduce stress peaks (see Section 3.2) do not offer sufficient solutions to this problem (Kollmann, 1984; Kröger et al., 2018; Mather et al., 1972). However, complex adjustments of the shaft outer or hub inner contour can achieve the goal of eliminating undesirable variations of the pressure within the joint and/or the failure-critical principal stress. However, these iterative numerical approaches of (Blacha, 2009) and (Krautter et al., 2015) reach such complexity that even experts refuse to accept them as an economic technique for optimization. In particular, the reason for this is that a local adjustment of the contour also affects adjacent areas, resulting in complex interactions.

Using AI as a solution approach for contour optimization, which, in addition to the previous optimization of the pressure within the joint, also allows direct optimization of the failure-critical principal stress, thus providing significant added value for the user. Especially, this added value is characterized by an increased efficiency, which is achieved by automatization and by learning the relations of neighboring areas. Here, learning is performed by coupling numerically calculated stress results with the corresponding geometric change of the interstice. From the overall objective of this paper, the requirements for an AI approach are derived as follows:

- High degree of generalization of the optimization target, so that stress optimization can be applied in many ways
- Self-determined selection of the optimization parameter in the shrink fit by the operator (pressure within the joint, main stress or shear stress)
- Specification of a target variable profile along the interstice; tailored to the needs and requirements of the application
- No expert knowledge required, such as specification of limits for interferences, stresses, etc.

4. Artificial Intelligence

4.1. Basic considerations

In the field of Artificial Intelligence (AI), there are now many terms being used synonymously and sometimes mixed together or are not clearly differentiated from one another. The general term AI includes various other terms. Systems that are listed under the term AI basically serve to imitate human thought and action (European Parliament, 2020). This includes, for example, machine learning (ML), which is relevant to the optimization problem presented. According to Mitchell, ML is spoken of when

"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks in T , as measured by P , improves with experience E ." (Mitchell, 1997)

Subcategories of ML are Supervised Learning (SL), Unsupervised Learning (UL) and Reinforcement Learning (RL) (see Figure 5, left) (Skansi, 2018). The subdivision takes place depending on the data situation and / or the optimization goal. If a computer independently learns complicated contexts by mapping them and by combining simple contexts, the result is a multi-layered model that shows the interconnection of many contexts in artificial neural networks. This form of information processing is called Deep Learning (DL) (Goodfellow et al., 2016). (Figure 5) (left) illustrates the categorization of terms within AI. SL requires a training data set that clearly links inputs to outputs. From the point of view of the corresponding algorithms, the relationship between input and output is thus known from the beginning. The goal in this case is to find the best possible generalization for the dataset so that future events can be predicted in the best possible way based on the training dataset (Cunningham et al., 2008). An example of supervised learning would be weather forecasting based on a large amount of historical data. UL, on the other hand, is used with a training data set that has only inputs (Ayodele, 2010). The goal of the corresponding unsupervised learning algorithms is to independently find previously unknown patterns in the data and thus establish unique relationships. An example of this would be the derivation of categories in messages (IBM, 2020). RL is an agent-based algorithm, where the agent independently develops a strategy to maximize the value of the reward function. The value of the reward function is used to directly evaluate the agent's action, thereby approximating a utility function from which the agent adjusts its strategy. The method of DL can be applied to any method of ML and also to its hybrid forms (Skansi, 2018). DL is suitable for processing large data sets in combination with artificial neural networks. This last is a data structure that is modelled like the human brain. Based on the basics of the different machine learning methods and their possible applications presented above, a suitable method can be selected in the following, together with the requirements from section 3.4. UL is the least suitable for the use case presented here, since this method is particularly suited for identifying clusters in data sets and thus dependencies between several parameters (Ayodele, 2010). RL and SL are better suited, since they can be used to describe the relationship between input and output reliably (Nasteski, 2017). Furthermore, RL differs positively from SL because the agent-based method does not need to be provided with initial training data (Sutton et al., 2018). Thus, RL is also suitable for application by non-experts. For the reasons explained above, RL is the most suitable approach to find an optimal solution in a fully automated way, without the necessity of having expert knowledge or initial training data. Therefore, the AI-based method for optimization presented in this paper is RL.

4.2. Definition of Reinforcement Learning

In RL, the process is determined by the agent-environment interaction (see Figure 5, right). The agent executes the action a_t at a time t based on the reward r_t and the state s_t . This action influences the environment in the downstream time step $t+1$ in such a way that the agent can change the reward r_t . This action influences the environment in the subsequent time step $t+1$ in such a way that the agent receives the reward r_{t+1} for its action a_t and is now confronted with the state s_{t+1} , which in turn leads to the cycle starting again from the beginning by a renewed action execution of action a_{t+1} . The policy π_t based on the states governs the probability of selecting a particular action a_t from the range of all possible actions A . (Sutton et al., 2018)

The **policy** π_t describes the procedure of an agent to assign a suitable action to a state at a certain point in time. It is essential for the model quality and can be very simple or highly complex. The **reward signal** represents the core of an RL process. After each iteration step, the agent is rewarded with a number (reward signal). The reward signal helps the agent to distinguish between good and bad actions. The agent always pursues the goal of maximizing his reward. The **value function** describes the sum of all expected rewards for a state. It is thus a long-term orientation for the agent and gives an indication of which chosen actions could lead to which states. The **model** allows the RL algorithm to predict subsequent actions and states from a combination of action and state. A model is thus needed for the prediction of states before they have even occurred. (Sutton et al., 2018)

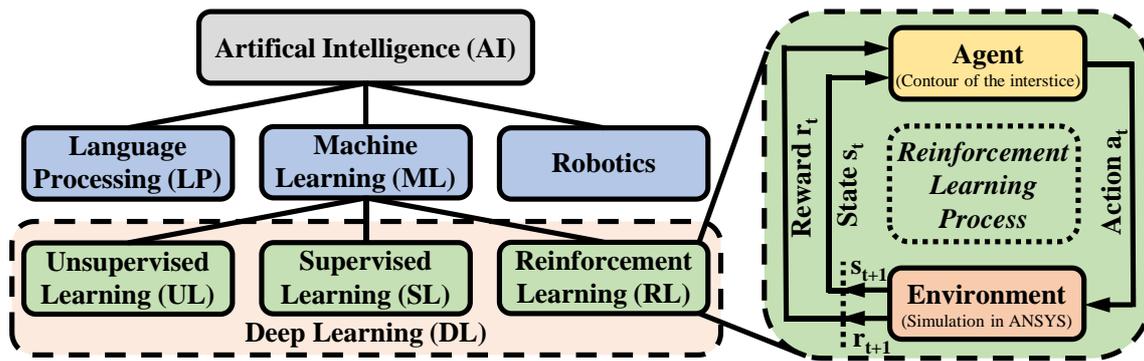


Figure 5. Schematic classification of the most important terminology within AI (left) and schematic illustration of the process of an RL algorithm (right)

5. Integration of AI into stress optimization on shrink fits

The implementation of RL into the FEM-based optimization of a shrink fit offers the option that profiles of stresses along the interstice can be generated by contouring. In this paper, the AI-based contour optimization is applied to the shaft. In principle there is nothing to prevent application of the same procedure to the hub. If an RL algorithm from (Figure 5) (right) is transferred to the usage case presented here, the simulation of the shrink fit corresponds to the environment. Here, the shrink fit, i.e. the environment, is discretized in the area of the interstice by a number n of transversely displaceable points on the shaft (cf. Figure 6). These n points represent the function of the agent (cf. Figure 5, right). By the possible actions A , in the form of a transversal displacement of the agent, the interference at any point can thus be increased, decreased or maintained. The states S represent the deviation determined between the plane stress condition σ_{psc} and the numerical simulation σ_{FE} , and results from the shape of the radial point positions (cf. Figure 6). This results in the reward of a point according to Equation 5:

$$r_n = 1 - \left[\left(\frac{|\sigma_{FE,n} - \sigma_{psc}|}{\sigma_{psc}} \right) \right] = 1 - \left(\frac{|\Delta\sigma_n|}{\sigma_{psc}} \right) \quad (5)$$

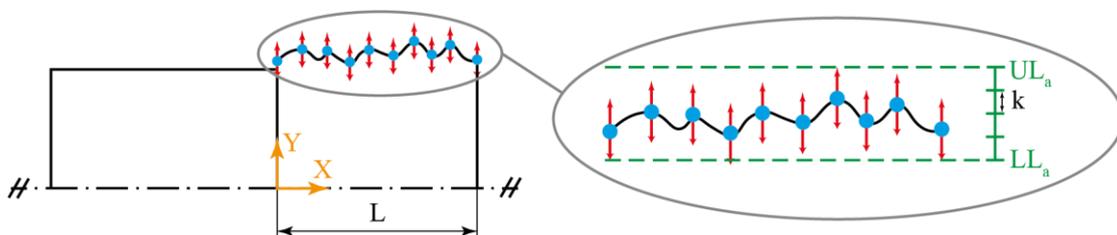


Figure 6. Interstice discretized by means of n transversely movable points

A common algorithm within reinforcement learning is the so-called Q-learning. Here, a table, the Q-table, is created. This table is updated continuously. The table contains the possible actions A in columns and the possible states S in rows. The values of the table correspond to the Q-values (see Figure 7, left). The table thus acts as an action specification for the agent. For each of the n points a state vector s contains i state descriptions (Equation 6). This results in a matrix of $n \cdot i$ possible states S (Equation 7).

$$s_{1 \dots n} = [\Delta\sigma_{n,1} \dots \Delta\sigma_{n,i}]^T \quad (6)$$

$$S = [s_1 \dots s_n]^T \quad (7)$$

The determination of i is done via Equation 8 by setting an upper limit UL_s and a lower limit LL_s for the state s and choosing a state step size j :

$$i = \frac{UL_s - LL_s}{j} \quad (8)$$

Applied to the shrink fit discussed here, a state vector s is the description of the possible stress deviation at a point. The lower and upper state limits are determined on the basis of σ_{psc} . The determination of the

step size j corresponds to the achievable accuracy of the homogenization and is also determined by σ_{psc} . If either of the two state limits UL_s or LL_s is exceeded, no further state can be assigned from the table. In this case, a directional correction is to be provided at the affected point(s) (maximum interference change, which requires an increase or decrease of the interference depending on the situation). This must be repeated until a state is reached once again in the Q-table. The possible actions A are the same at all points and correspond to the transverse displacement in discrete form. The required upper and lower action limits (UL_a and LL_a , respectively) result from the interference U_{psc} of the plane stress condition. The step size k is determined based on the manufacturing process (predominantly CNC turning). This results in the possible number of interference changes m according to Equation 9:

$$m = \frac{UL_a - LL_a}{k} \quad (9)$$

From this follows for action a (Equation 10):

$$a_{1..m} = \Delta U_{1..m} \quad (10)$$

and for the vector A to (Equation 11):

$$A = [\Delta U_1 \dots \Delta U_m] \quad (11)$$

A tabular overview of all values and correlations for the shrink fit can be found in Table 1. The structure of the Q-Table is shown schematically in (Figure 7) (left).

Table 1. Overview of defined values

Symbol	Description	Defined as	Symbol	Description	Defined as
UL_s	Upper state limit	$1,5 \cdot \sigma_{psc}$	UL_a	Upper action limit	$1,5 \cdot U_{psc}$
LL_s	Lower state limit	$0,5 \cdot \sigma_{psc}$	LL_a	Lower action limit	$0,5 \cdot U_{psc}$
j	State step size	$0,05 \cdot \sigma_{psc}$	k	Action step size	$1 \mu m$

Each combination of a state s to an action a can be assigned a Q-value. This Q-value allows the agent to receive a valuation for his action a . According to Bellman's optimality principle (Equation 12), the quality value Q is a function of s and a at a given time t (Sutton et al., 2018):

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \cdot \left[r_{t+1} + \gamma \cdot \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right] \quad (12)$$

This equation contains the current Q-value $Q(s_t, a_t)$, the learning rate α , the reward r_{t+1} for the action a_t , the discount rate γ and the maximum expected future Q-value for action a in the next state s_{t+1} . The learning rate α influences the rate at which the agent uses learned knowledge to modify the existing Q-value. At a lower Q-value the algorithm learns more slowly and is therefore more stable, whereas at a higher Q-value it mainly considers recently learned knowledge. The discount rate γ can be used to differentiate between a high reward in the immediate and a more far-sighted strategy. The goal is to identify the maximum possible Q-value for each of the n points. For each Q-value $Q(s, a)$, the action a is associated with the state s , resulting in an optimal contour. At the beginning of Q-learning, all Q-values are initialized to 0. This poses a problem, because the algorithm cannot use the maximum Q-value to decide which action is the best possible. In Q-learning, there is a danger that a local optimum will be reached too quickly, because the process is increasingly focused on the Q-values (greedy strategy) so that the global optimum cannot be reached. These considerations show the importance of the trade-off between exploration and exploitation. Q-learning can explore the environment through exploration and thus discover new areas. However, the search to achieve an optimum is not triggered systematically but occurs as the result of random variations. Through exploitation, Q-learning always strives towards the currently known optimum. With the help of an epsilon-greedy action selection, Q-learning can achieve a trade-off between exploration and exploitation. Epsilon (ϵ) stands for the probability with which the algorithm does not select the best possible action according to the Q-values ($a_{t, greedy}$) but instead randomly selects another possible action ($a_{t, \epsilon-greedy}$). Thus, several optima can be found. Epsilon does not necessarily have to be a constant. There are approaches which adapt epsilon during learning (Tokic, 2010). An epsilon-greedy strategy is used for the present optimization problem.

The optimization process described above was finally applied on cylindrical shrink fits with simple geometries similar to (Figure 1Figure 2d). It became apparent that an oscillating stress curve initially results in the early exploration phase (cf. Figure 7, center). This behavior of the algorithm can be justified by the previously chosen epsilon-greedy strategy, which explores the solution space in the first iterations. However, with increasing number of iterations, the algorithm has gathered experience and the epsilon-greedy strategy changes towards exploitation. Therefore, an improvement of the stress profile can be observed (cf. Figure 7, center). In order to validate the gained results, a proven indirect method (Ulrich et al., 2016; Kröger et al., 2018; Kröger et al., 2020) was chosen, since there is no direct method to measure the pressure within the interstice. This indirect method compares the diameter expansion of the hub within the experiment and the numerical calculation. For this purpose, the specimens were measured before and after joining on a coordinate measuring machine, so that the resulting diameter expansion could be compared with the numerically calculated values. Such a comparison also requires a precise representation of the materials used, which was ensured before by precision tensile tests. The results show a very high accordance (Figure 7, right), which is why they are considered validated.

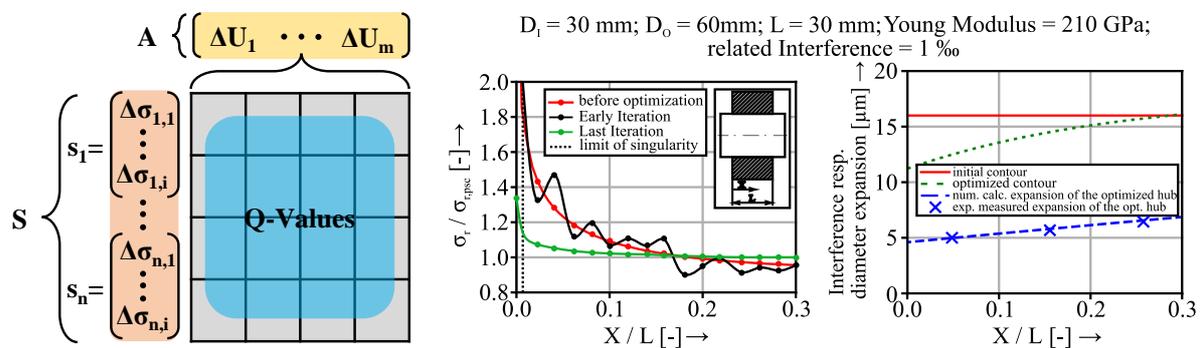


Figure 7. Qualitative representation of the Q-table for the presented use case (left), initial application on a sample shrink fit (center) and generated contour as well as validation (right)

6. Conclusion and Outlook

Increases in the power density of drive systems inevitably result in increased stresses within shrink fits. These can cause the permissible stresses of the materials to be reached or exceeded, in particular at stress peaks (cf. Section 3.1). Simple geometric optimization options already exist (Leidich, 1983; Kollmann, 1984), but these suffer from various disadvantages (Kröger et al., 2018) (cf. Section 3.2). The approaches developed at the research center (Blacha, 2009; Krautter et al., 2015) permit achievement of almost constant pressure profiles within the joint. However, application of these approaches is very complex, time-consuming, imprecise and also requires expert knowledge. Furthermore, the approaches are so far suitable only for optimizing the pressure within the joint (cf. Section 3.3).

The RL-based optimization method presented in Section 5 is capable of addressing these remaining drawbacks. In terms of complexity, the user's work is reduced by the fact that the algorithm developed can be easily transferred to any geometries of shrink fits. Therefore, the user only has to define the geometry to be optimized and the desired interference. Additionally, this is why expert knowledge is no longer required. Moreover, instead of the pressure within the joint, any stress such as the tangential stress σ_t , which is critical for the failure of the shrink fit, can be used as an optimization parameter, resulting in an enhanced material utilization. Finally, the epsilon-greedy Q-learning algorithm is able to find the optimal solution to a problem independently.

Currently, the research center is working on an optimization of the algorithm to avoid a contouring of the interstice which is not manufacturable. This will increase the performance of the algorithm in terms of saving time, since these implausible solutions will not be considered from the beginning. In future, the results obtained with the AI support presented in Section 5 will be made available in an (online) database, so that where solutions to problems already exist these can be accessed directly or by interpolation, without the need to perform a new numerical simulation. This will gradually reduce the

expert knowledge required to find solutions, allowing any designer to perform such optimization with little effort. Furthermore, future Q-learning can be extended to incorporate a neural network, which takes us into the territory of deep-Q-learning. In that particular usage case, the neural network can help find the optimal action within a continuous solution space, thus further improving the optimization results. Beyond this usage case, the optimization algorithm can be adapted for many other applications.

References

- Ayodele, T. O. (2010), "Machine Learning Overview", In: Zhang, Y. (Ed.), "New Advances in Machine Learning", IntechOpen, Rijeka, pp. 9-18. <https://doi.org/10.5772/225>
- Blacha, M. (2009), "Grundlagen zur Berechnung und Gestaltung von Querpressverbänden mit Naben aus monolithischer Keramik", PhD thesis, University of Stuttgart.
- Cunningham, P., Cord, M., Delany, S. J. (2008), "Supervised Learning", In: Cord, M., Cunningham, P. (Eds.), "Machine Learning Techniques for Multimedia: Case Studies on Organization and Retrieval", Springer, Berlin, pp. 21-49. https://doi.org/10.1007/978-3-540-75171-7_2
- DIN (2017), DIN 7190-1:2017-02, "Pressverbände - Teil 1: Berechnungsgrundlagen und Gestaltungsregeln für zylindrische Pressverbände", Beuth Verlag GmbH, Berlin.
- European Parliament (2020), "Was ist künstliche Intelligenz und wie wird sie genutzt?", [online] <https://www.europarl.europa.eu/news/de/headlines/society/20200827STO85804/was-ist-kunstliche-intelligenz-und-wie-wird-sie-genutzt> [retrieved on 09.11.2021].
- Goodfellow, I.; Bengio, Y. & Courville, A. (2016), "Deep Learning", MIT Press.
- Glögler, C. (2003), "Untersuchungen an spannungshomogenisierten und zylindrischen Pressverbindungen unter Torsionsbelastung", PhD thesis, University of Stuttgart.
- Gropp, H., Ziaei, M. (2012), "Tendenzielle Ermittlung von zulässigen Werten für das erweiterte Ruiz-Chen-Kriterium bei reibdauerbeanspruchten torsionsbelasteten Pressverbindungen", VDI-Berichte vol. 2176, Düsseldorf, pp. 37-47.
- Hartmann, M. (1999), "Konstruieren mit Keramik: Ohne Ecken und Kanten", Industrieanzeiger Nr. 5.
- IBM (2020), "Unsupervised Learning", [online] <https://www.ibm.com/cloud/learn/unsupervised-learning> [retrieved on 09.11.2021]
- Kollmann, F. G. (1984), "Welle-Nabe-Verbindungen: Gestaltung, Auslegung, Auswahl", Springer, Berlin. <https://doi.org/10.1007/978-3-642-61727-0>
- Krautter, M., Binz, H. (2015), "Improvement Of The Designing Method Of Hybrid Interference Fits" NAFEMS World Congress, 2nd International SPDM Conference, San Diego.
- Kröger, J., Binz, H. (2018), "Spannungsoptimierung von Pressverbänden mit additiv gefertigten Naben: Numerische und experimentelle Untersuchungen", VDI-Berichte vol. 2337, Stuttgart, pp. 199-210.
- Kröger, J., Binz, H. (2020), "Hohe Übermaße bei Pressverbindungen: Untersuchungen zu Auslegungsgrenzen und Steigerung der maximalen Übermaße bei zylindrischen Pressverbindungen", FVA Research Booklet 1399 and related information sheet, Frankfurt.
- Leidich, E. (1983), "Beanspruchung von Pressverbindungen im elastischen Bereich und Auslegung gegen Dauerbruch", PhD thesis, TH Darmstadt.
- Mather, J., Baines, B. H. (1972), "Distribution of stress in axially symmetric shrink-fit assemblies", Wear 21, pp. 339-360.
- Mitchell, T. M. (1997), "Machine Learning", McGraw-Hill, New York.
- Nasteski, V. (2017), "An overview of the supervised machine learning methods", HORIZONS. B. 4, pp. 51-62.
- Rentzsch, W. H., Willmann, G. (2002), "Ein einfaches Hilfsmittel zum Konstruieren mit Keramik", In: "Materials Science & Engineering Technology", vol. 33, pp. 184-189.
- Skansi, S. (2018), "Introduction to Deep Learning: From Logical Calculus to Artificial Intelligence", In: Mackie, I. (Ed.), "Undergraduate Topics in Computer Science", Springer. <https://doi.org/10.1007/978-3-319-73004-2>
- Sutton, R. S. & Barto, A. G. (2018), "Reinforcement Learning: An Introduction", The MIT Press.
- Tokic, M. (2010) "Adaptive ϵ -Greedy Exploration in Reinforcement Learning Based on Value Differences", In: Dillmann, R., Beyerer, J., Hanebeck, U.D. and Schultz, T. (Eds.), KI 2010: Advances in Artificial Intelligence, 2010, Berlin, Heidelberg, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 203-210. https://doi.org/10.1007/978-3-642-16111-7_23
- Ulrich, D., Binz, H. (2016), "Einfluss von Schmierstoffen aus der Massivumformtechnik auf die Reibdauerbeanspruchung mikroschlupfanfälliger Welle-Nabe-Verbindungen: Numerische und experimentelle Untersuchungen anhand zylindrischer Querpressverbände mit beschichteten Wellen unter wechselnder Torsionslast", VDI-Berichte vol. 2287, Karlsruhe, pp. 65-78.
- Vidner, J., (2016) "Methode zur Bewertung der Ermüdungsfestigkeit von reibdauerbeanspruchten Systemen", PhD thesis, TU Chemnitz.