



**Országos Tudományos Diákköri
Konferencia
2022**



**Iterative Identifiability Analysis of
Composite Material Failure Models**

Készítette: Ipkovich Ádám

Témavezető: Dr. Abonyi János, Dr. Kummer Alex

Mérnöki Kar, Folyamatmérnöki Tanszék

A pályamunka lezárásának dátuma: 2022.11.06

Contents

1	Introduction	2
2	Theoretical background of the identifiability analysis and iterative experimental design algorithm	5
2.1	Formulation of the identification problem	5
2.2	Experimental design-based identifiability analysis	9
2.3	Linear programming-based optimal experiment selection – iterative design of experiments	11
2.4	The proposed iterative parameter identification algorithm	12
3	Results and discussion	15
3.1	Model selection	15
3.2	Model analysis	17
3.3	Iterative design of experiments	18
4	Conclusion	22
A	Supporting material on the failure models	23
A.1	Description of the studied models	23
A.2	Failure surfaces of the use case	27

Abstract

Composite materials are becoming more significant in the industry due to their strength-to-weight ratio. The behavior of the plies differs by the materials used and can be described by failure models. A failure model is only identifiable if stress in the direction of and parallel to the fiber direction and the shear stress is measured, and the parameters of the models provide us with hints about when the material breaks. Yet, various parameters are interconnected, and most identification processes require lots of data extracted from time-consuming and expensive fracture tests. My work focuses on the identifiability of the failure models, in other words, the parameters of the failure surface. The first step focuses on selecting the failure model by identifying the parameters and calculating the errors of the models. The best model is carried over to determine the parameters. The Fisher Information Matrix is an excellent tool for extracting information from the experiments efficiently from even a small number of correct experiments. The method calculates the partial derivatives (local sensitivity) of the failure model and determines which experiment is relevant to which parameter. As the exact number of required experiments is unknown and the various experiments may not provide additional information, I use a Design of Experiments approach to select the ones that contain the most information. Several optimal designs exist, from which I use the A-optimality due to its additive nature, and thus the optimality measure presents the information contained in a singular value. This measurement is optimized through a non-linear mixed integer optimization to select only the relevant experiment. The algorithm can also determine which parameter requires additional information and propose optimal points on the failure surface. Lastly, an iterative algorithm is introduced to select the points that contribute the most to the identification process. I introduce the effectiveness of my method by identifying well-known failure models' parameters on random-generated failure surfaces. The identification process includes the Max-Stress, Tsai-Wu, Tsai-Hill, Hoffman, Hashin, and Puck models. The results show that the method provides parameters close to the nominal parameters of the generated failure surface and proposes tests if no information is available.

Keywords: Fisher Information Matrix; Model identification; Composite failure models; Design of experiments; Iterative parameter optimization

1 Introduction

Industrial sectors are becoming ever-more penetrated by composite materials. The increasing popularity can be credited to their strength and weight ratio. Consequently, the topic gained ground, and an immense effort was put into the behavior and failure of composite material products. Yet, some composites exert anisotropy and complex fracturing mechanisms with various modes, which increased the difficulty of composite material failure analysis [1]. Failure and strength analysis is a fundamental step in the design process to determine the risk of failure and provide guidance on mitigating its occurrence. Some examples include wind turbine strength analysis [2], composite joints modeling [3], 3D simulation of metal bearings failure [4], plywood bending [5], glass fiber drilling [6], *etc.*

The behavior of the composite materials is modeled with several orthotropic failure models, most of which are often available in commercial Finite Element Method (FEM) software packages. The most popular ones are the Max Stress [7], Tsai-Wu [8], Hoffman [9], Tsai-Hill [10], Hashin [11], and Puck [12] failure criteria. In our research, I investigated classical failure models for a shell-like composite structure, represented by orthotropic constitutive behavior with translaminar (*e.g.* fiber cleavage, buckling, and shearing) and intralaminar failure modes (through-thickness failures).

Machine learning has already been used in composite model identification. The Tsai-Wu model was employed to simulate unidirectional composite lamina, which was optimized by a neural network [13]. Soft computing has also been used in material fatigue estimation [14]. A Micromechanics-based approach for Static Failure (M-SaF) and Bayesian statistical framework has been proposed to predict the failure in heterogeneous materials [15]. The identification of the model parameters is performed using Bayesian Inference method. There is some work in the literature where global optimization algorithms (*e.g.* genetic algorithms) are used to determine the model parameters, but no work can be found where the identification task is supported by experimental design. Despite the fact, that Fisher Information Matrix (FIM) has been successfully applied in parameter optimization in many cases [16], the parameter estimation of material failure models with the design of experiments is unknown in the literature, hence I propose to use FIM-based optimal design to support the model identification problem.

The identification of the composite material of the model composite material is not straightforward. The most common practice is to perform experiments, provoking characteristic stress-strain states, exploring the failure characteristic along a few radial trajectories of the failure envelope and use these test data to calibrate numerical prediction models. Accurate identification requires informative experiments to contribution to describing material failure parameters. Different models are expressed by different parameter sets that can be characterized with dissimilar experiments. Inaccurate test data inevitably adds

error and uncertainty, mainly from imperfect knowledge of the material behavior, indirect phase space (stress space), scatter, and noise in the experimental data [17]. Therefore, our goal is to a) find the right failure model with the best possible fitting quality under the given data set and b) find the most informative material tests for the model fitting. This can lead to a minimum test matrix to be defined for any pre-selected failure model that provides the most information about all parameters and thus, results in robust and reliable model fit including uncertainty quantification and survival probability dependent design.

For a robust and fast algorithm development, I used virtual tests to generate the data set to develop the methodology. It gave us the opportunity to directly validate the methodology by knowing the real sought parameters upfront. Moreover, extending the nominal virtual tests with systematic and unsystematic scatter could give us valuable feedback about the behavior of the algorithm and quantify its robustness. However, in a future work, the algorithm should be tested using real test data.

Identifiability analysis is closely related to sensitivity analysis. If the objective function does not vary with a model parameter, it follows that the parameter in question cannot be readily estimated from the available tests [18]. The set of experiments must contain information about all the model parameters; otherwise, some of the model parameters may not be identifiable, leading to a false model prediction. To reduce the number of experiments and retain accuracy, I propose the application of experimental design [19]. It allows us not only to consider identifiability but also to perform the most informative experiments possible. The advantage is that only a set of predefined (minimum) experiments are required, improving the time and cost-efficiency of data-driven modeling. The use of the optimal design of experiments (DoE) is recommended as not only the selection of the most informative experiments but ignoring bad/inconsistent or redundant data is also possible with its help [20].

In this work, the identifiability analysis is based on the trace of the FIM, performed in local points of the ply failure stress state, along with optimal designs criteria (*i.e.* A-, E- and D-optimality) [21]. The most informative experiments are chosen based on their A-optimality, as the contribution of all experiments can be linearly decomposed into the contribution of the individual experiments. The key idea is to accurately identify the model that best describes the material with the least amount of data. The best model is selected based on how the measured data fits on the failure surface generated by the identified parameters, and the best one is carried over to the identification process.

FIM-based sensitivity analysis is proposed to investigate the identifiability of the parameters. I propose an iterative experimental design and parameter fitting method to identify model parameters. In each iteration, the experiments are selected based on their

contribution to identifying the model parameters, and the iteration stops if the model performance improvement becomes insignificant. Moreover, if the characteristic behavior of the composite fly failure models is required for a certain material, the experimental data points must contain information on all parameters. However, this is seldom the case in the industry environment, and the available data set should define which model can be calibrated best for the specific application.

Various test data sets have been tested with the models. However, a particular model may dominate the other in the same material with a different condition. For example, the stack-up design, amount of layers, and their orientation may alter the governing failure model [22]. Therefore, there is a need to select the model that best fits the material behavior and its nominal values. In most cases, the model parameters are determined by specific experiments, *i.e.* the aim is to determine a model parameter as precisely as possible. However, some parameters cannot be measured, and an optimization algorithm is needed.

Hence, our work provides the following contribution:

- I present a compact methodology to decide which composite failure model should be used for simulation based on the available experimental data. As such, the failure models can be systematically analyzed and evaluated in terms of model accuracy and model parameter identifiability (see section 2);
- I evaluate the performance of the failure models to select the best one based on the Euclidean distance of the experimental points from the failure surface estimated by data reconciliation (see subsection 2.1);
- I performed a Fisher Information Matrix-based sensitivity analysis to investigate the identifiability of model parameters using experimental data (see subsection 2.2);
- I prove the performance of the proposed method using a procedurally generated experimental data set. The visualization of the sensitive regions of the failure surface is presented with the recommended and most informative experiments for parameter identification. The feasibility of the parameter identification algorithm is demonstrated in section 3.

The structure of the paper is as follows: The theoretical background for the iterative identifiability analysis of composite material failure models are introduced in section 2. Subsection 2.1 describes the generalized form of the failure models, the experiments and parameters, whereas subsection 2.2 includes the tools, such as the Fisher Information Matrix and optimal design. The selection problem is described in subsection 2.3, and the algorithm is detailed in subsection 2.4. The results are presented in section 3, where both

the analysis of the models and the iterative DoE are presented on procedurally generated data. In section 4, concluding remarks are given. An Appendix is also provided for the mathematical description of failure models (Appendix A.1) and their visualization (Appendix A.2).

2 Theoretical background of the identifiability analysis and iterative experimental design algorithm

This chapter introduces the methodology for the identifiability analysis of composite materials. First, the identification problem is formalized in subsection 2.1 that is followed by the theoretical background of optimal selection (subsection 2.2), including the sensitivity, Fisher Information Matrix and optimal design. The selection problem is defined in subsection 2.3, before the algorithm is described in subsection 2.4.

2.1 Formulation of the identification problem

The general formula of failure materials describes material breakage with failure stresses and specific parameters. The goal of the models is to approximate material failure through parallel, perpendicular (to fiber direction), and shear stress. Damage occurs when the failure model value reaches one, calculated from the model if failure tests and specific parameters present accurate values. As various failure models are available, the particular models are denoted by f_m , where m represents the index of a failure model. In addition, each model bolsters a set of parameters, some of which are unique to a specific physical phenomenon. The general notation for a failure model is as follows:

$$f_m(\mathbf{x}, \theta^{(m)}) = 1 \quad (1)$$

$$\mathbf{x} = [\sigma_1, \sigma_2, \tau_{12}] \quad (2)$$

where $\theta^{(m)}$ is considered the parameter vector of the failure model f_m , which takes up the value one when the material would break; therefore, $f_m(\mathbf{x}, \theta^{(m)}) = 1$ denotes the marginal state of the failure, the fracture of fiber or matrix in the composite material. $\mathbf{x}$ represents a failure in the stress state space that consists of three variables: unidirectional tension/compression parallel (σ_1) and perpendicular to (σ_2) the fiber direction, and shear stress (τ_{12}).

The parameters are imperative to predict breakage correctly and are often unknown.

A model is identified through the optimization of its parameters, which can be formalized as follows:

$$\theta^{(m)} = [\theta_1^{(m)}, \theta_2^{(m)}, \dots, \theta_n^{(m)}] \quad (3)$$

where $\theta^{(m)}$ defines the parameter vector with length n .

The Tsai-Wu model is considered one of the most popular failure criteria [8]. The model consists of a single continuous function and six parameters. In the case of the Tsai-Wu failure model, the parameter vector is defined as:

$$\theta^{(\text{Tsai-Wu})} = [X_t, X_c, Y_t, Y_c, S_{12}, F_{12}] \quad (4)$$

where $\theta^{(\text{Tsai-Wu})}$ denotes the parameter vector of the Tsai-Wu model, X_t and X_c stand for the tension and compression in the direction of the fibers, respectively, Y_t and Y_c describe the tension and compression perpendicular to the fiber direction, whereas S_{12} is the shear failure parameter. F_{12} is a parameter unique to the Tsai-Wu model.

As such the model is formulated as in Eq. 2.1 with its parameters defined in Eq 4:

$$f_{\text{Tsai-Wu}}(\mathbf{x}_k, \theta^{(\text{Tsai-Wu})}) = 1 = \frac{\tau_{12}^2}{S_{12}^2} - \sigma_2 \left(\frac{1}{Y_c} - \frac{1}{Y_t} \right) - \sigma_1 \left(\frac{1}{X_c} - \frac{1}{X_t} \right) + \frac{\sigma_1^2}{X_c X_t} + \frac{\sigma_2^2}{Y_c Y_t} + \frac{2 F_{12} \sigma_1 \sigma_2}{\sqrt{X_c X_t Y_c Y_t}} \quad (5)$$

All other models incorporate $[X_t, X_c, Y_t, Y_c, S_{12}]$ parameters, the coefficients of the failure for the axes of the stress space. Some models, such as the Puck, Hashin, and Tsai-Wu failure criterion, also contain unique parameters. The Max Stress, Hoffman, Hashin, Tsai-Hill, and Puck failure models are described in Appendix A.1.

Model parameters require fine-tuning to identify the highest accuracy model. The goal is to estimate the parameters as accurately as possible. N_p number of experiments is available that contains as much information as possible while keeping the number at a minimum so that the parameters are efficiently identified. One realization of failure (experiment) takes up the form of $\mathbf{x}_k$ that consists of the principal stresses $[\sigma_{1,k}, \sigma_{2,k}, \tau_{12,k}]$. The experiments are organized into a matrix $\mathbf{X}^{(N_p)}$, and the optimized parameters are identified based on the information content of the available tests. As such, the model errors stemming from the parameters are minimized based on their difference to the breakage status ($f_m = 1$):

$$\mathbf{X}^{(N_p)} = [\mathbf{x}_1^T, \dots, \mathbf{x}_{N_p}^T] \quad (6)$$

$$\epsilon = \frac{1}{N_p} \sum_{k=1}^{N_p} (f_m(\mathbf{x}_k, \theta^{(m)}) - 1)^2; \quad k = 1, \dots, N_p \quad (7)$$

$$\theta^{(m,0)} = \arg \min_{\theta^{(m)}} \epsilon \quad (8)$$

where $\theta^{(m,0)}$ denotes the optimized parameters of the model, $\mathbf{x}_k$ stands for the k th experiment $\mathbf{x}_k = [\sigma_{1,k}, \sigma_{2,k}, \tau_{12,k}]$ from matrix $\mathbf{X}^{(N_p)}$, with $k = 1, \dots, N_p$ denoting the running index of the experiments. ϵ defines the mean squared error of the predicted model failure to breakage.

Although the mean squared error of a model provides information on how well it fits, these error values cannot be compared between models as various modes denote different functionalities in behavior. Therefore, a robust measure is required to evaluate the performance of the models. The quality of the fitting can be evaluated by calculating the Euclidean distances of the experiments and their corresponding value during breakage returned by the nominal failure model. The values of the experiments and projected ones do differ: most test results contain measurement errors that heavily depend on the actual state of the material. The projection is also known as data reconciliation [24], the goal of which is to calculate the Euclidean distance of the experiments to the closest point on the model while simultaneously projecting the data. The method resembles the least squares method [25]:

$$\begin{aligned} \arg \min_{\hat{\mathbf{x}}_k} e_k &= (\mathbf{x}_k - \hat{\mathbf{x}}_k)^T \mathbf{V} (\mathbf{x}_k - \hat{\mathbf{x}}_k) \\ \text{s.t.} \quad &f_m(\hat{\mathbf{x}}_k, \theta^{(m,0)}) = 1 \end{aligned} \quad (9)$$

where e_k is considered as the distance between the experiments and the closest failure model value, $\mathbf{x}_k$ denotes the stress state on failure, $\hat{\mathbf{x}}_k$ is considered as a projected point on the m th model with the nominal parameters $\theta^{(m,0)}$ and $\mathbf{V}$ describes the weight matrix. The function minimizes the distance value of the experiments to the nominal failure surface, whose median indicates the performance of the identified model.

Additionally, the argmin calculates the closest projected point to the experiment. Figure 1 depicts the individual Euclidean distance error (e_k) of the k th experiment. The Euclidean distance between the experiments and the identified failure model is minimized, while simultaneously, the experiment is reconciled to the optimized failure model and added in the set of errors $E^{(m)}$, whose median defines the fitness of the nominal failure.

Performing failure tests are time-consuming and costly, and the number of available

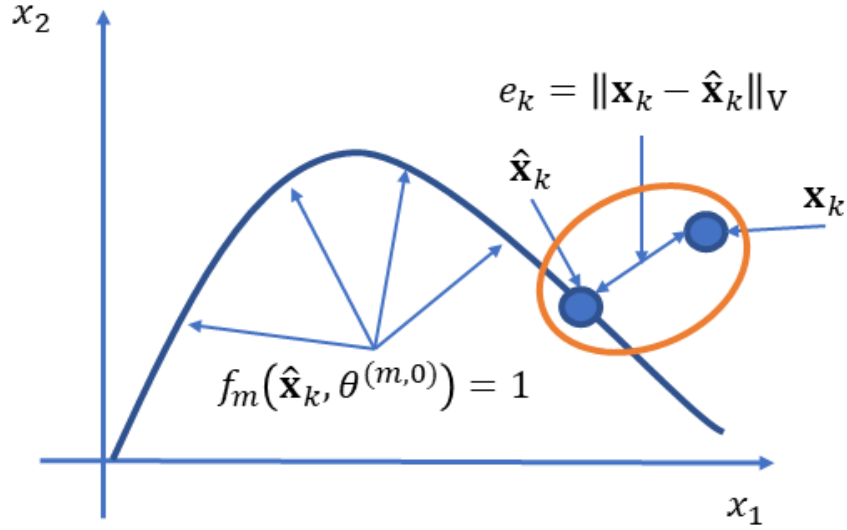


Figure 1: **Data reconciliation-based evaluation of the failure surfaces.** The fitness of failure surface (f_m) is evaluated by measuring distances between the original experiment $\mathbf{x}_k$ and the closest experiments $\hat{\mathbf{x}}_k$ on the nominal failure model ($f_m(\hat{\mathbf{x}}_k, \theta^{(m,0)})$).

experiments is often insufficient. Yet, if enough points are selected to identify the parameters, the variable stress space can be reproduced and examined to propose the necessary experiments that may help in the identification process. Thus, a failure surface is generated to present which part of the failure surface contributes the most to parameter identification. The potential experiments ($\mathbf{X}^{(N_s)}$) with N_s amount of experiments are sampled from the failure surface, which also contains the optimal experiments ($\mathbf{X}^{(N_p)}$). The number of potential experiments must be one order of magnitude more than the optimal amount ($N_s \gg N_p$). The failure surface may be generated with a grid-like allocation of two principal stresses $\mathbf{x}'_k$, in this case, the $[\sigma_1, \sigma_2]$ variable pairs. The last one is optimized so that the substitution into the model results in a one ($f_m(\mathbf{x}_k, \theta^{(m)} = 1)$). As an alternative, the Latin Hypercube [27] can be used to sample data by dividing the probability function into equal parts and selecting points from the partitions equally. The optimizer finds the corresponding τ_{12} value from the generated stress pairs. In this work, the grid-like approach is applied to generate the failure surface:

$$\arg \min x_k^* \in \mathbf{x}_k \left(f_m([\mathbf{x}'_k, x_k^*], \theta^{(m,0)}) - 1 \right); k = 1, \dots, N_s \quad (10)$$

where x_k^* scalar denotes the optimized stress, where as the vector $\mathbf{x}'_k$ denotes the given stress pairs. Although the optimized scalar may be any of the principal stresses, this work optimizes the $x_k^* = \tau_{12,k}$ shear stress.

There are several advantages of generating a failure surface. It is possible to:

- Determine the most informative experiments without having to conduct more failure tests.

- Visualize the failure of the composite material at any eligible state.
- Calculate the information content of the potential experiments for the entire surface.

Numerous theoretical composite ply failure models have been developed based on different continuum mechanical assumptions that are not always consistent. The failure surfaces heavily depend on the parameter identification process. Moreover, the experiments include errors that change with the dataset, so two datasets of the same material may provide different nominal parameters. As such, there is a need to measure how an experiment may contribute to the identification of a model. Performing sensitivity analysis and experiment design help describe the information content of the failure surfaces.

2.2 Experimental design-based identifiability analysis

Any experiment may provide a certain contribution to the identifiability of the parameters. As the models are based on physical behavior, each parameter has its particular meaning *e.g.* the Tsai-Hill model has five elementary parameters that can be best identifiable according to their stress axis. Each experiment has a contribution to the parameter identification process, but that value depends on where the point is located in the stress space. Some parameters require more than one stress axis (*e.g.* the Tsai-Wu coefficient), and a group of parameters may be simultaneously identifiable with only one specific experiment. One should calculate the contribution of each parameter so that the set of optimal solutions can be determined.

The local sensitivity $\mathbf{s}_k$ of the k th experiment in the stress state space is an excellent measure for the contribution of the point, which can be calculated by partially differentiating the model by the parameter vector:

$$\mathbf{s}_k = \left. \frac{\partial f_m(\mathbf{x}_k, \theta)}{\partial \theta} \right|_{\theta=\theta^{(m,0)}} = \left[\frac{\partial f_m(\mathbf{x}_k, \theta)}{\partial \theta_1}, \dots, \frac{\partial f_m(\mathbf{x}_k, \theta)}{\partial \theta_n} \right]^T; k = 1, \dots, N_p \quad (11)$$

where $\mathbf{s}_k$ denotes the sensitivity of the parameters regarding the k th experiment.

Also, the sensitivities of the experiments can be organized into a matrix:

$$\mathbf{S}^{(N_p)} = [\mathbf{s}_1, \dots, \mathbf{s}_{N_p}] \quad (12)$$

where $\mathbf{S}^{(N_p)}$ stands for the sensitivity matrix of $\mathbf{X}^{N_p}$ optimal experiments.

Although the sensitivities provide information separately on the contributions to the parameters, standard tools are available to calculate their aggregated information content. Aggregation is performed with optimal design. The Fisher Information Matrix (FIM)

functions as the basic structure for the aggregation: it presents the possible information content of a set of experiments for both individual, and joint sensitivities [21]. The FIM of the k th experiment ($\hat{\mathbf{F}}_k(\theta^{(m)})$) is calculated by the matrix product of the parameter sensitivity vector found in Eq 11, a user-defined orthogonal matrix $\mathbf{W}_{n \times n}$ that defines which parameters are the priority, and the transpose of sensitivity vector. The FIM ($\hat{\mathbf{F}}(\theta^{(m)})$) of the set of experiments is considered the sum of the individual FIMs.

$$\begin{aligned} \hat{\mathbf{F}}(\theta^{(m)}) &= \sum_{k=1}^{N_p} \hat{\mathbf{F}}_k(\theta^{(m)}) = \sum_{k=1}^{N_p} \mathbf{s}_k^T \mathbf{W} \mathbf{s}_k = \\ &= \sum_{k=1}^{N_p} \left. \frac{\partial f_m(\mathbf{x}_k, \theta^{(m)})}{\partial \theta^{(m)}} \right|_{\theta^{(m)}=\theta^{(m,0)}}^T \mathbf{W} \left. \frac{\partial f_m(\mathbf{x}_k, \theta^{(m)})}{\partial \theta^{(m)}} \right|_{\theta^{(m)}=\theta^{(m,0)}} \end{aligned} \quad (13)$$

where $\hat{\mathbf{F}}(\theta^{(m)})$ denotes the Fisher Information Matrix of a set, $\hat{\mathbf{F}}_k(\theta^{(m)})$ stands for the Fisher Information of the k th experiment, $\mathbf{s}_k$ describes the sensitivity of an experiment, while $\mathbf{W}$ represents an orthogonal weight matrix. The priority of model parameters can be varied so that more information is extracted by modifying the weight values of the $\mathbf{W}$ matrix. In this work, equal priorities are used. However, weights can be chosen to put a stronger emphasis on model parameters where uncertainty is higher.

Optimal design finds the most informative points in the set of available experiments. Each design is connected to minimizing the variance of the parameters by optimizing the FIM. Several approaches exist, such as the E-, D-, and A-optimality; each requires the FIM matrix to qualify the experiment. E-optimality maximizes the minimum eigenvalues of the FIM, while D-optimality maximizes the determinant. A-optimality maximizes the sum of the diagonal. The advantage of A-optimality is that it can be decomposed into the individual A-optimality values of the experiments. As such, the individual contribution of the experiments can be calculated as a simple addition:

$$\hat{A} = \text{tr}(\hat{\mathbf{F}}(\theta^{(m)})) = \sum_{k=1}^{N_p} \sum_{i=1}^n s_{k,i}^2 \quad (14)$$

where $\hat{A}$ denotes the A-optimality of the set of experiments, k denotes the index of the experiment, i stands for the index of the parameter. The total A-optimality value is calculated as the trace of the FIM, in other words, the sum of the diagonal.

If the parameters, the models, and two of the stress variable vectors are known (or generated), it is possible to calculate the information content of the failure surface and determine its sensitivity concerning the parameters. Also, the theoretical maximum of the parameter sensitivities can be defined, which can be used to normalize the sensitivity

values. By understanding how the experiments contribute to the identification process, I can select the optimal ones and recommend them to determine and provide pointers on the behavior of the material.

2.3 Linear programming-based optimal experiment selection – iterative design of experiments

In order to determine the parameters properly, an N_p amount of optimal experiments should be chosen from the N_s potential experiments. As newly added experiments may not contain additional information, the goodness of an individual point must be established so that only the necessary optimal ones are selected. The A-optimality provides the required information on the contribution of the experiments; thus, the remaining problem is to select the ones that contribute the most to each parameter. One of the advantages of the Fisher Information Matrix is that the contributions to the identification of each model parameter can be defined and investigated individually. However, A-optimality has an advantage over the FIM: the parameter identification can be solved as a linear programming problem due to the additive nature of the A-optimality. Note that the identifiability of the parameters must be ensured with a predefined amount of experiments. Therefore, I assume that the set of potential experiments N_s contains N_p optimal experiments that provide enough information to identify the parameters.

The selection of the most informative points requires an optimization algorithm, which is a nonlinear mixed-integer solver [26]. A binary integer y_k is assigned to each experiment determining whether the experiment is chosen. N_p represents the maximum number of experiments to be selected, so the sum of y_k must not exceed the total number of experiments. This way, it is ensured that only the N_p most informative points are selected. The optimization problem itself is to maximize the total A-optimality ($\sum_{k=1}^p y_k \sum_{i=1}^n u_{k,i} s_{k,i}^2$) of the optimal set of experiments, while the squared sensitivity associated with the i th parameter of the chosen points is more than or equal to a b_i constraint value. Moreover, the sensitivity values can be weighted based on the importance of the parameter (or its accuracy), $u_{k,i}$ can be used to represent the weight of the k th experiments by the i th parameter.

$$\begin{aligned} & \max \sum_{k=1}^N y_k \sum_{i=1}^n u_{k,i} s_{k,i}^2 \\ \text{s.t. } & \sum_{k=1}^N y_k u_{k,i} s_{k,i}^2 \geq b_i \quad \forall i \in 1, \dots, n \text{ and } y_k \in \{0, 1\} \text{ and } \sum_{k=1}^N y_k \leq N_p \end{aligned} \quad (15)$$

where y_k represents whether the k th experiment is selected, $s_{k,i}^2$ stands for the squared sensitivity vector of the parameters that is compared to a predefined b_i constraint. N_p

represents the number of feasible experiments to select from the set of experiments. $u_{k,i}$ represents the weight of the i th parameter sensitivity of the k th experiment.

The optimization determines which experiments have a significant impact on parameter identification. The design of experiments is expected to be performed in the appropriate variable space, where the k th experiment contains information to identify at least one parameter. It is also possible to tweak the algorithm to propose mechanical experiments. By selecting the theoretical point with the maximum sensitivity constrained to one of the principal stresses, it can be associated with single-axis failure tests.

2.4 The proposed iterative parameter identification algorithm

This section describes the iterative identification process. Figure 2 and Algorithm 1 provide visual and textual drafts of the algorithm; the content of both is discussed in the following paragraphs.

The inputs of the algorithm include a starting set of failure tests $\mathbf{X}^{(0)}$ and the model - parameter pairs ($F = \{(f_1, \theta^{(1)}) \dots, (f_M, \theta^{(M)})\}$) to extract the optimized (nominal) parameter $\theta^{(m,0)}$. The algorithm currently contains six failure models, namely the Max Stress [7], Tsai-Wu [8], Hoffman [9], Tsai-Hill [10], Hashin [11], and Puck [12] failure criteria. The input data is reconciled $\hat{\mathbf{X}}^{(m)}$ to evaluate how well the identified failure model fits the data. A failure model f_m is considered the best if the median of the set $E^{(m)}$ of distances between the experimental data and failure models is the least, and all its parameters are identifiable. The best model is referred to as f_b , and its nominal parameters are denoted by $\theta^{(b,0)}$.

In order to select the optimal experiments, the failure surface is generated to provide the potential experiments $\mathbf{X}^{N_s}$. N_s is sampled, from which the sensitivity matrix $\mathbf{S}^{(N_s)}$ is calculated to examine which experiment contributes to a particular parameter. To select the optimal points $\mathbf{X}^{(N_p)}$, I use an iterative design of experiments approach to select only the most informative points. It is important to note that, assuming that each parameter is independent of the other, N_p needs to be at least the length of the parameter vector (n) as each parameter requires at least one experiment with high sensitivity (regarding that parameter) to be chosen. In some cases, where the parameters are interconnected, the N_p number of experiments may be decreased.

The DoE part of the algorithm begins by selecting the optimal points. The iteration starts with $j = 1$, where j is the number of iterations. During the selection of the optimal points, the constraint vector $\mathbf{b}$ requires comparable values, so the sensitivities are normalized by the theoretical maximum of the failure surface. Then, the optimization problem described in Eq. 15 is solved with $\mathbf{U}$ weight matrix to prioritize parameters (if required). After the optimal points are selected, the parameters are reidentified, from

which the A-optimality $\hat{A}_j$ is calculated by summing the diagonal of the FIM of the selected experiments. Lastly, the mean squared error ϵ is calculated from the reidentified models and the selected experiments. Stopping the iteration requires selecting all points or the error to be less than a threshold value t . The A-optimality and error values of the selected experiments are always recalculated. The parameters constantly improve over each iteration, and the selection order defines the importance of the new set of experiments. It is also possible to modify the algorithm to only select the necessary experiments, and stop. The output returns the best model (f_b), the selected experiments ($\mathbf{X}^{(N_p)}$), the corresponding parameters ($\theta^{(b,j)}$) and their A-optimality ($\hat{A}_j$).

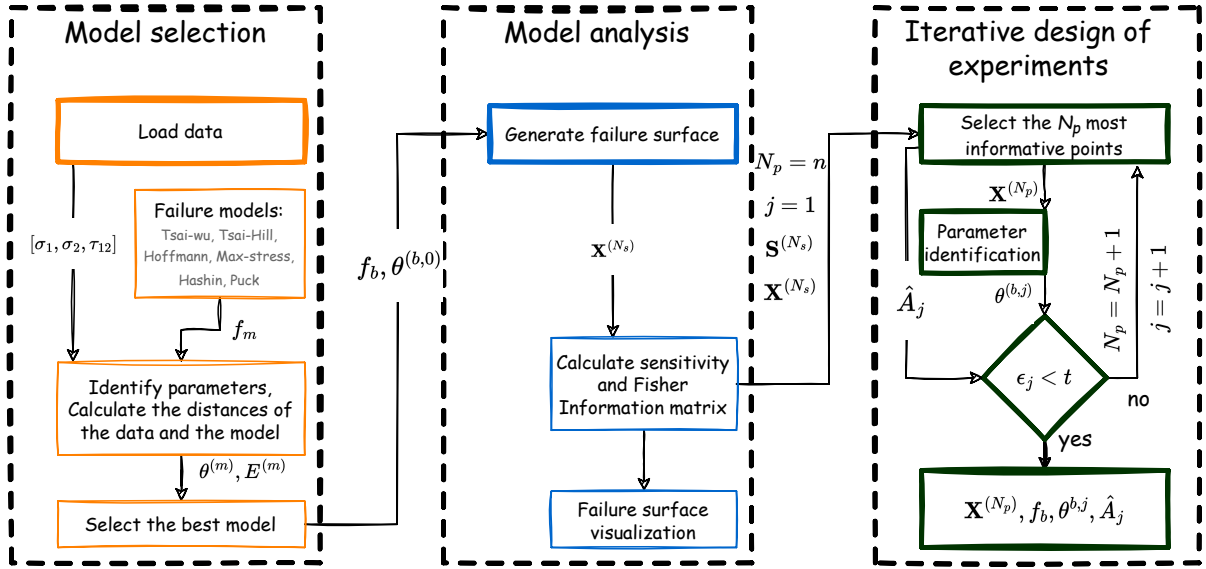


Figure 2: **Flowchart of the iterative identifiability algorithm.** The algorithm requires the data and the failure models to be loaded and forwarded to the model selection function. The chosen model has the best fit for the data. The parameters are identified, and a failure surface is generated before the sensitivity of the experiments is calculated. The Fisher Information matrix of the individual experiments is determined so that the N_p most informative points can be selected through the fitting error and A-optimality ($\hat{A}_j$). If the results are inadequate, N_p is incremented by one, and the parameters, MSE (ϵ_j), and A-optimality are recalculated regarding the added optimal experiment. If adequate, the algorithm stops.

Algorithm 1 Description of iterative DoE-based experiment selection.

Input: $\mathbf{X}^{(0)}$, F ▷ $\mathbf{X}^{(0)}$ denotes the set of experiments, $\mathbf{x}_k = [\sigma_{1,k}, \sigma_{2,k}, \tau_{12,k}]$; $k = 1 \dots N_0$.

▷ The failure model set F contains the M amount of failure models f_m and their starting parameters $\theta^{(m,0)}$

for $m = 1, \dots, M$ **do**

$\theta^{(m)} = \text{ParameterIdentification}(f_m, \mathbf{X}, \theta^{(m,0)})$ ▷ Identify the parameters from a number of starting experiments.

$\hat{\mathbf{X}}^{(m)} = \text{ReconcileDataToFailureSurface}(f_m, \mathbf{X}, \theta^{(m)})$ ▷ Reconcile $(\hat{\mathbf{X}}^{(m)})$ the starting experiments $\mathbf{X}^{(0)}$ so that they fit on the

failure surface $f_m(\cdot, \theta^{(m)})$.

$E^{(m)} = \text{GetDistance}(\mathbf{X}, \mathbf{V}, \hat{\mathbf{X}}^{(m)})$

▷ Calculates the distances of all experiments $\mathbf{X}^{(0)}$ from the nominal surface $\hat{\mathbf{X}}^{(m)}$, and returns the set of the distances $E^{(m)}$.

▷ $\hat{\mathbf{X}}^{(m)}$ stands for the reconciled data for failure model f_m

end for

$f_b, \theta^{(b,0)} = \text{SelectBestModel}([\theta^{(1)}, \dots, \theta^{(M)}], [E^{(1)}, \dots, E^{(M)}])$

▷ Returns the best performing model and its parameters based on the the median of the Euclidean distances

$\mathbf{X}^{(N_s)} = \text{GenerateFailureSurface}(\theta^{(b,0)}, \text{resolution})$ ▷ Generates the failure surface of the model with $\theta^{(b,0)}$ parameters.

▷ The resolution of the surface is also required (e.g. 10×10 grid)

$\mathbf{S}^{(N_s)} = \text{CalculateSensitivity}(\mathbf{X}^{(N_s)}, \theta^{(b)})$ ▷ Calculates the sensitivity of the experiments (individually) of the sampled surface $\mathbf{X}^{(N_s)}$.

$N_p = n$

▷ The starting experiment number equals to the number of parameters so that each has at least one experiment to be identified from.

$j = 1$ ▷ Defines the running index for the while cycle.

while $\epsilon_j < t$ and $N_p \leq N_s$ **do** ▷ t stands for a user-defined threshold value.

$\mathbf{X}^{(N_p)} = \text{SelectOptimalPoints}(\mathbf{S}^{(N_s)}, \mathbf{U}, \mathbf{b}, N_p)$ ▷ Calculates the FIM, and selects optimal points, where N_p number of experiments

are selected.

▷ The nonlinear mixed integer optimizer is balanced by a constraint vector $\mathbf{b}$. The total contribution of the parameters must exceed the corresponding constraint.

$\theta^{(b,j)} = \text{ParameterIdentification}(f_b(), \mathbf{X}^{(N_p)}, \theta^{(b,j-1)})$

▷ The model based on the optimal points is reevaluated by identifying the parameters and recalculating errors.

▷ $j = 0$ is defined as the starting parameters $\theta^{(b,start)}$.

$\hat{A}_j = \text{tr}(\hat{\mathbf{F}}(\theta^{(b,j)}))$ ▷ Calculating the A-optimality

$\epsilon_j = \text{MSE}(f_b, \theta^{(b,j)}, \mathbf{X}^{(N_p)})$ ▷ Calculate the mean squared error of the optimal points.

$j = j + 1$

$N_p = N_p + 1$

end while ▷ Returns the best model, optimal points, parameters and the A-optimality values

Output: $f_b(), \mathbf{X}^{(N_p)}, \theta^{(b,j)}, \hat{A}_j$

3 Results and discussion

The performance of the described algorithm is illustrated through a procedurally generated dataset with 5% standard deviation errors using the Tsai-Hill failure model. The nominal parameters were $[X_t = 2134, X_c = 541, Y_t = 87.5, Y_c = 214, S_{12} = 62.4]$. A total of fifty marginal states of the failure in the σ_1 - σ_2 - τ_{12} stress space were sampled from all over the failure surface, so it is ensured that all parameters are identifiable. The fifty experiments act as the starting matrix for the identification process. Due to the low noise, I expect that the identifiability analysis will determine the model correctly and identify the parameters with a small error. The models include the Max-Stress, Tsai-Wu, Hoffmann, Tsai-Hill UD, Hashin, and Puck failure criteria. Moreover, I recommend experiments by manually highlighting the most sensitive uniaxial experiments on the failure surface by picking those whose stress value is only dominant in one axis *e.g.* $\mathbf{x}_k = [2000, 0, 0]$.

I have programmed the algorithm in Python. Therefore, several different Python packages have been used (*e.g.* pandas and NumPy) [28]. The Euclidean distance error (data reconciliation) and optimization algorithms were implemented with scipy/lmfit optimizer [29, 30]. As the model selection is a simple ranking, I calculate the medians with pandas and NumPy. The generation of the failure surface uses the NumPy linspace function for a 2D grid of $[\sigma_1, \sigma_2]$, and τ_{12} is optimized through scipy/lmfit with the help of the calculated parameters, implemented failure models and the generates stress values. The partial differentiation for establishing the sensitivity matrix is provided by the numdifftools package [31]. The selection of optimal experiments happens through the SCIP mixed-integer nonlinear solver from the ortools package [32].

3.1 Model selection

If the actual parameters and governing principles behind the behavior of the material is unknown, the best-fitting models should be selected first. The evaluation process requires the optimal parameters and uses the minimum Euclidean distance of the measured point from the optimal failure surface. The median of the distances provides information on which model performs the best with a given set of experiments. Figure 3 presents the boxplots of the Euclidean distances shown on the y -axis. Most models (located on the x -axis) perform comparably well due to their similar failure surfaces. The Tsai-Wu, Tsai-Hill, and Hoffman models are half ellipsoids but somewhat different in their orientation. Appendix A.2 depicts the shapes of the failure surfaces. It is important to note that while the previously mentioned failure models consistently provide ellipsoids, the Puck is incredibly flexible in its shape due to the four additional parameters. One can also configure it to be an ellipsoid. Meanwhile, the Max-stress is considered a cuboid and

can very well cover the stress space where the experiments can be located. As such, Max-stress may not be the best choice; if the failure model is an ellipsoid, then there is a possibility that the max-stress does cover non-populated parts of the stress space. The Hashin failure surface takes up a curved pyramid-like shape (simply put) and performs the worst, as the experimental data was sampled from an ellipsoid with fairly low noise.

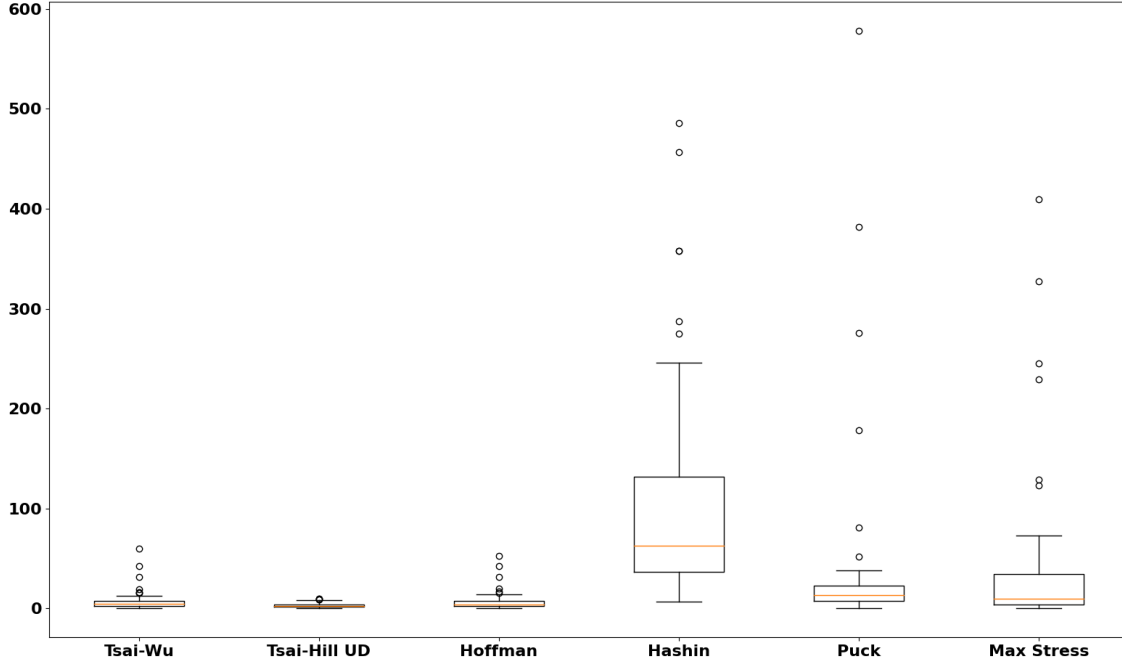


Figure 3: **Performance boxplots of composite failure models.** The Tsai-Hill model perfomed the best, while the Hashin and the Max Stress were the worst.

The best model is chosen based on how well the data fits on the model: the median of the Euclidean distances to the failure surface is considered. In this case, the Tsai-Hill criterion performs with the lowest mean distance, and Hashin and Max Stress is the worst. Most non-ellipsoid models have significant outliers that highlight that the models cannot describe various experiments and have blank spaces that cannot be fitted onto the data. The outcome is not surprising as the stress pairs were generated by the Tsai-Hill model, and the deviation of the data from the model is relatively low (5%). It is important to note that the Puck, Tsai-Wu, and Hoffman models perform quite well in terms of the median distances.

3.2 Model analysis

Following the model performance evaluation, I explore the parameter-sensitive stress regions of the Tsai-Hill failure surface. Therefore, I performed the model parameter sensitivity analysis and presented the sensitivity of the model parameters (X_t , X_c , Y_t , Y_c and S_{12}) on the failure surface in Figure 4.

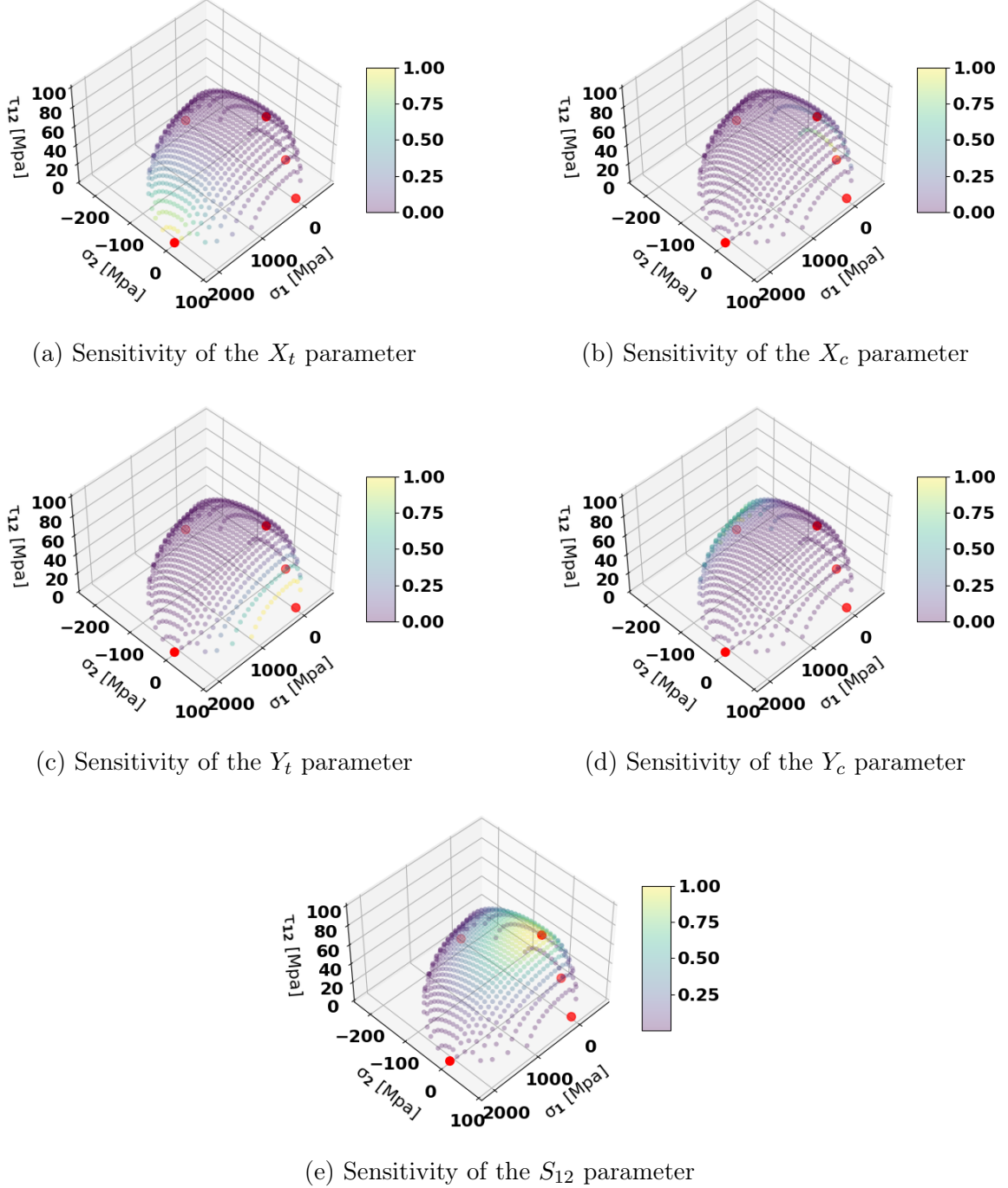


Figure 4: **Sensitivity of Tsai-Hill model parameters** X_t , X_c , Y_t , Y_c , S_{12} . The $[X_t, X_c, Y_t, Y_c]$ parameters are identifiable jointly, with only two experiments. S_{12} is connected to the shear force (τ_{12}), so it requires an experiment where the shear force is significant.

The color bar presents the sensitivity values on the failure surface. Each parameter sensitivity was normalized with the corresponding maximum value. The stress regions are shown for each investigated model parameter. Note that the sensitivities seem to be perpendicular to the related axis. The nominal parameter X_t presents where the material would break upon perpendicular tensile stress. The σ_1 has to be at least the nominal parameter to break the material. So to identify all experiments that may cause failure, an area should be drawn with the nominal parameter as the center and the uncertainty of the parameter as its width. This area is considered perpendicular to the stress axis as the σ_1 may not change significantly; on Figures 4, this area is the yellowish-greenish segment of the failure surface. This way, all the stress regions can be explored where the specific model parameters can be identified. Some points may contain information about more than one parameter, *e.g.* the subfigures contain points on the mesh that contributes to both X_c and Y_t . It is preferable to execute the experiments on the vertices of the failure mesh. Experimental design can be used to define theoretical stress pairs with the highest sensitivity to a model parameter.

3.3 Iterative design of experiments

The result of the iterative experimental design and parameter fitting is presented in Figure 5. The depicted values present how the A-optimality monotonously increases from iteration to iteration while the mean squared error (MSE) drops significantly. The MSE is insignificant (≈ 0.035) from the first iteration due to the low noise of the experimental data, yet, the MSE, in general, decreases further until the last iteration. However, based on the model accuracy of the values of MSE, there is no need to include additional experiments after the 34th iteration step as the model accuracy does not improve significantly. Although the A-optimality is monotonous, it may not necessarily be the case for the error value. If the uncertainty is high or too many uninformative experiments are selected, the MSE may begin to rise again. After some additional inclusions, it may also drop, so stopping the algorithm at the right time may be difficult for the analyst. Figure 5 shows how the new experimental points were added during the iterations. The blue dots presents the experimental points up to the iteration stop condition (so until 34th step), and the grey dots can be defined as not included.

Figure 6 presents the fitted model parameters in every iteration step, and the black vertical line presents the 34th step, where the model parameters are accepted as nominal parameters. As the experimental points were generated with the same model (Tsai-Hill model), the known nominal parameter can also be depicted in the figure, which is presented by a black horizontal line. As such, comparing the nominal and identified parameters shows that the algorithm determines the parameters in the vicinity of the

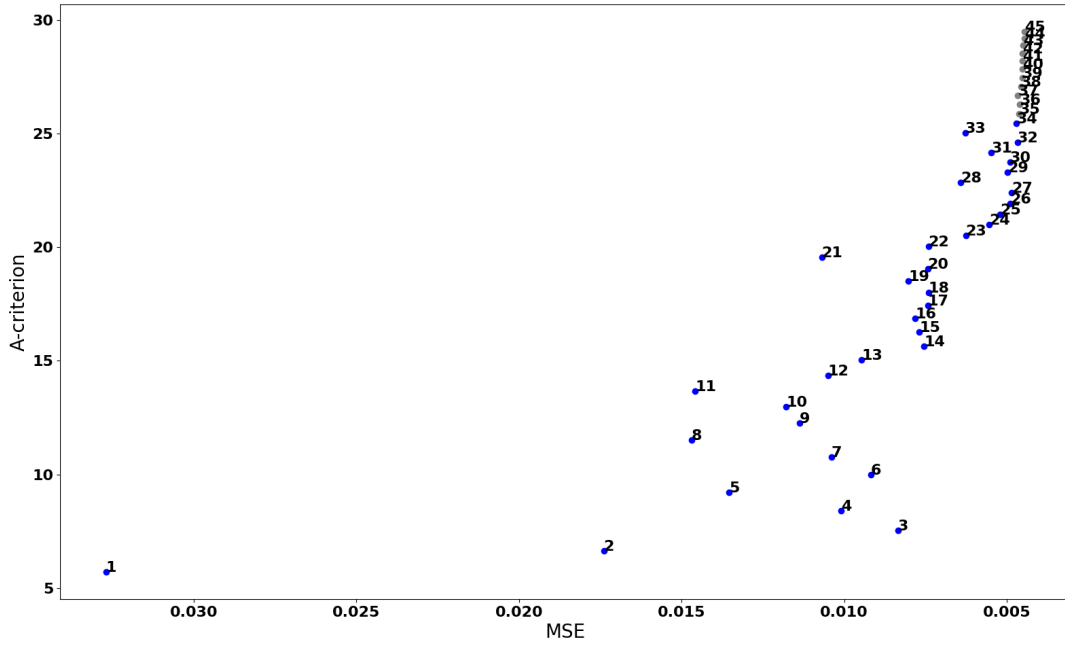


Figure 5: **MSE and A-optimality values during the iterative experiment design and parameter fitting.** While the A-optimality monotonously increases, the error may deviate. The second batch of experiments may not be selected as the errors do not improve significantly over many iterations. The point that are not selected are colored in grey, whereas the blue points are the ones that are chosen.

nominal parameters, with a minor difference.

The iterative experiment design algorithm aims to cover the entire stress field and collect information from most locations. The iterative DoE algorithm selects experiments in a manner that promotes obtaining the maximum information. The selection seems circular; in other words, the algorithm chooses a point far away from the last one and draws a circle if the dots of each iteration are connected. The selector considers the whole of the failure surface, not just one region, and in doing so, identifies the parameters with accuracy. However, the identification process may overfit the model. With the inclusion of non-informative experiments, the DoE starts to model the errors.

Parameters that fluctuate may be interconnected with other parameters, so the other parameters may react and change their values significantly. Moreover, the parameters may find a local extreme, so they fail to change significantly. If the algorithm stops at the right time, overfitting can be avoided.

Figure 7 presents the identified failure surface in the σ_1 - σ_2 - τ_{12} plane, where the blue dots denote the experimental data stress points that were considered in the parameter identification, and the black/grey dots represent that the experimental points were uninvolved. The identified Tsai-Hill surface describes the experimental data well. The

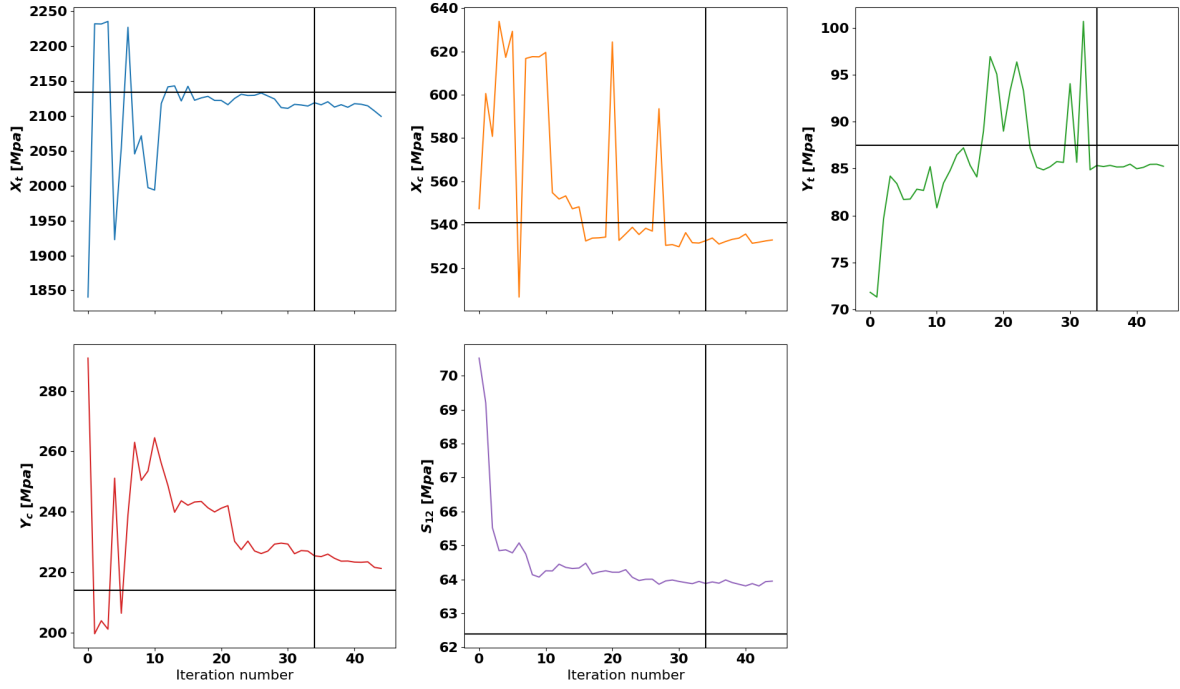


Figure 6: **Fitted model parameters in the iteration process.** This figure presents the fitted model parameters in every iteration step, and the black vertical line presents the 34th step, where the model parameters are accepted as nominal parameters. The horizontal black line presents the nominal model parameters used in experimental data generation.

black/gray dots are mostly the ones that significantly deviate from the failure surface. The blue ones seem to fit perfectly as the model describes their behavior rather than the ones with heavy noise. Therefore it is important to stop before the surface is fitted on the black points, which may significantly disturb the identification process and causes overfitting, *i.e.* the noise modeling.

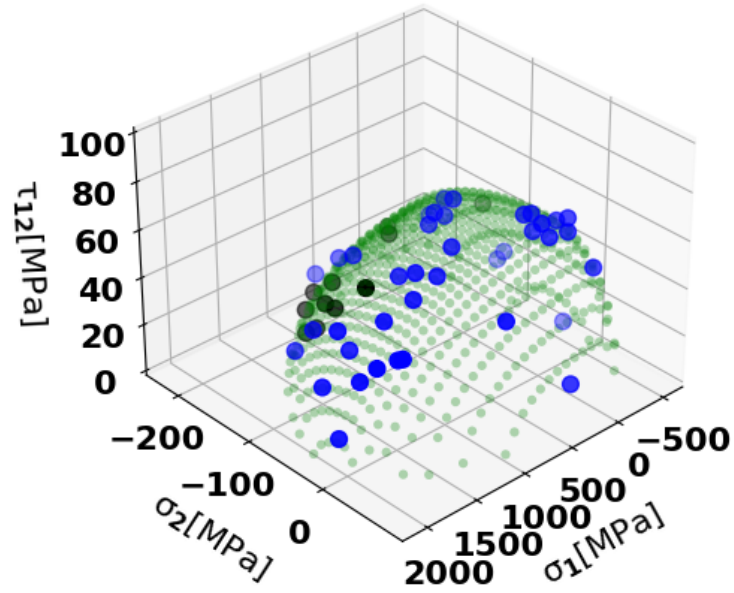


Figure 7: **Tsai-Hill failure surface with the generated experiments.** The surface is identified in the iteration process with a nominal set of parameters. The blue dots indicate the experimental points used for identification and the grey ones that were unused.

4 Conclusion

In this work, I proposed a methodology to identify the composite material failure models and select experiments. Material behavior may be described with different failure models, such as the Tsai-Wu, Max Stress, Hoffman, Hashin, Tsai-Hill, and Puck models. I proposed a robust method to evaluate which model best fits the experimental data through the Euclidean distances of the experiments and failure surfaces.

When the model is selected, the failure surface is generated to provide information on where the parameters are best identifiable. Unfortunately, this work does not include an algorithm to propose experiments as failure tests. However, the uniaxial experiments with the most information content are presented. The method requires little change to be able to suggest failure tests. Highlighting the uniaxial experiments has a disadvantage; some of the parameters of more complex models, such as the Puck or Hashin, may need to be identified from more than one stresses. Clustering the surface segments and selecting the experiments from the segments may be a good solution, so it is a future work of the author.

The results propose that the identification method can select the most informative experiments. It is important to note that the basis of the analysis, the Tsai-Hill model, is a relatively simple model as it contains only the elementary parameters $[X_c, X_t, Y_c, Y_t, S_{12}]$. Also, the data features little deviation from the original failure surface. As such, noisier and more complex data may reduce the accuracy of the method. Nevertheless, the robust evaluation based on the Euclidean distances between the experiments and the nominal failure model does show promising results in selecting the appropriate model. I would also note that the data may "change" failure surfaces if higher noise is applied, so when working with heavily noisy data, the method may propose a different model.

In the case of the Tsai-Hill model, the experiments-based approach design yielded satisfying results in model fitting. Although it is impossible to precisely determine the parameters of a composite material, the approximation of parameters for the generated data is accurate. The algorithm selects points from all over the failure surface. If some surface segments are underrepresented, then the method may not be able to determine the parameters. It is a future work of the author to perform proposed methodology on physical experiments.

The iterative identifiability analysis can select experiments to determine the model parameters accurately; for composite materials, it may help reduce the cost and time efficiency of modeling the behavior.

Acknowledgments

This research was funded by the Foundation For the University of Pannonia (Scholarship for the Scientific Student's Association Conference: PA-S2R-0089/2022). Also, I would like to give my deepest appreciation to my supervisors, János Abonyi who encouraged me to give my best to complete this work and spared incomprehensible amount of time to help my efforts come true. Words cannot express my gratitude to Alex Kummer for his exceptional patience and immense work in helping me write this work.

A Supporting material on the failure models

A.1 Description of the studied models

Max-Stress criterion is one of the simplest composite failure model [7]. The interactions between the three stress components are not taken into account, but a distinction between the compressive and tensile loading of the material is considered. The criterion consists of five independent states, and five model parameters ($\theta = [X_t, X_c, Y_t, Y_c, S_{12}]$). X_t and X_c denote the tensile and compressive strength parameters in the direction of the fibers, respectively. Y_t and Y_c stand for tensile and compressive strength parameters perpendicular to the direction of fibers, S_{12} represents the in-plane shear strength parameter of the material, and

$$f_{\sigma_1} = \begin{cases} \frac{\sigma_1}{X_t} & \text{if } \sigma_1 \geq 0 \\ \frac{|\sigma_1|}{X_c} & \text{if } \sigma_1 < 0 \end{cases} \quad f_{\sigma_2} = \begin{cases} \frac{\sigma_2}{Y_t} & \text{if } \sigma_2 \geq 0 \\ \frac{|\sigma_2|}{Y_c} & \text{if } \sigma_2 < 0 \end{cases} \quad f_{\tau_{12}} = \frac{|\tau_{12}|}{S_{12}} \quad (16)$$

A material failure occurs if any of the three function ($f_{\sigma_1}, f_{\sigma_2}, f_{\tau_{12}}$) reaches value one.

$$f = \max(f_{\sigma_1}, f_{\sigma_2}, f_{\tau_{12}}) \quad (17)$$

Tsai-Wu model is considered one of the most popular failure criteria [8]. The model consists of a single continuous function and six parameters ($\theta = [X_t, X_c, Y_t, Y_c, S_{12}, F_{12}]$). F_{12} is considered as a coefficient of the Tsai-Wu model. Each model require uni-directional parallel (σ_1) and perpendicular (σ_2) compression/tension and shear stress (τ_{12})

The Tsai-Wu model can be described as:

$$f = \frac{\tau_{12}^2}{S_{12}^2} - \sigma_2 \left(\frac{1}{Y_c} - \frac{1}{Y_t} \right) - \sigma_1 \left(\frac{1}{X_c} - \frac{1}{X_t} \right) + \frac{\sigma_1^2}{X_c X_t} + \frac{\sigma_2^2}{Y_c Y_t} + \frac{2 F_{12} \sigma_1 \sigma_2}{\sqrt{X_c X_t Y_c Y_t}} \quad (18)$$

where f denotes failure of the material that theoretically takes up the value of one.

The Tsai-Wu model can be further simplified [33]; the **Hoffman criterion** is similarly quadratic polynomial in nature as Tsai-Wu is. The Hoffman failure criterion is almost the equivalent of the Tsai-Wu failure criterion with the exception of the disappearing Tsai-Wu coefficient as is shown in Equation 19, being equivalent to Tsai-Wu model with $F_{12} = 0$. Hoffman-criterion consists of a continuous function and five parameters ($\theta = [X_t, X_c, Y_t, Y_c, S_{12}]$).

$$f = \frac{\tau_{12}^2}{S_{12}^2} + \sigma_2 \left(\frac{1}{Y_t} - \frac{1}{Y_c} \right) + \sigma_1 \left(\frac{1}{X_t} - \frac{1}{X_c} \right) + \frac{\sigma_1^2}{X_c X_t} + \frac{\sigma_2^2}{Y_c Y_t} - \frac{\sigma_1 \sigma_2}{X_c X_t} \quad (19)$$

where f denotes failure of the material that theoretically takes up the value of one.

In case of the **Tsai-Hill criterion** two independent functions describe the failure, depending on the type of the reinforcement structure (unidirectional or fabric type). Unidirectional composites are materials whose layers consists of parallel fibres. The Tsai-Hill failure criterion for unidirectional composites is as follows:

$$f = \left(\frac{\sigma_1}{X} \right)^2 + \left(\frac{\sigma_2}{Y} \right)^2 + \left(\frac{\tau_{12}}{S_{12}} \right)^2 - \frac{\sigma_1 \sigma_2}{X^2} \quad (20)$$

where X denotes the strength of the material along the fiber direction, Y stands for the strength of the composite perpendicular to the fiber direction, S is considered the shear strength.

In order to incorporate differing tensile and compressive strengths X and Y are considered based on the values of σ_1 and σ_2 .

$$X = \begin{cases} X_t & \text{if } \sigma_1 \geq 0 \\ X_c & \text{if } \sigma_1 < 0 \end{cases} \quad Y = \begin{cases} Y_t & \text{if } \sigma_2 \geq 0 \\ Y_c & \text{if } \sigma_2 < 0 \end{cases} \quad (21)$$

Hashin-model considers fiber and matrix as different entities (f_f and f_m), and each has two discontinuous functions that describe the onset of failure. The maximum of their failure value decide when the material breaks. Hashin-model considers six parameters ($\theta = [X_t, X_c, Y_t, Y_c, S_{12}, Q]$). In fibers, the failure behaves as presented in Equation 22.

$$f_f = \begin{cases} \left(\frac{\sigma_1}{X_t} \right)^2 + \left(\frac{\tau_{12}}{S_{12}} \right)^2 & \text{if } \sigma_1 \geq 0 \\ \left(\frac{\sigma_1}{X_c} \right)^2 & \text{if } \sigma_1 < 0 \end{cases} \quad (22)$$

where f_f denotes the failure of the fiber of the composite. Whereas in the matrix, the

material failure can be described by Equation 23.

$$f_m = \begin{cases} \left(\frac{\sigma_2}{Y_t}\right)^2 + \left(\frac{\tau_{12}}{S_{12}}\right)^2 & \text{if } \sigma_2 > 0 \\ \frac{1}{Y_c} \left(\left(\frac{Y_c}{2Q}\right)^2 - 1 \right) \sigma_2 + \left(\frac{\sigma_2}{2Q}\right)^2 + \left(\frac{\tau_{12}}{S_{12}}\right)^2 & \text{if } \sigma_2 < 0 \end{cases} \quad (23)$$

where f_m denotes failure in the matrix of the composite, and Q stands for interlaminar shear stress. When either the fibres or the matrix fails, the material is considered broken. Therefore, the use of maximum is recommended:

$$f = \max(f_f, f_m) \quad (24)$$

The **Puck criterion** is one of the most complex models, consists of five different models describing the axial fiber failure (f_{ff}) and inter-fiber failure f_m . The simplified representation of fiber failure is the following:

$$f_{ff} := \begin{cases} \frac{\sigma_1}{X_t} & \text{if } \sigma_1 \geq 0 \\ \left| \frac{\sigma_1}{X_c} \right| & \text{if } \sigma_1 < 0 \end{cases} \quad (25)$$

A fracture occurs if f equals one in case of f_{ff} or if it equals weakening factor (f_w) in case of f_m . The weakening factor is calculated as:

$$f_w = \begin{cases} \sqrt{1 - c(|\sigma_1| - sX)^2} & \text{if } |\sigma_1| > sX \\ 1 & \text{if } |\sigma_1| \leq sX \end{cases} \quad (26)$$

$$c = \frac{1 - m^2}{[(1 - s)X]^2} \quad \text{where } 0 \leq m \leq 1, 0 \leq s < 1 \quad (27)$$

$$X = X_t \quad \text{if } \sigma_1 \geq 0 \quad (28)$$

$$X = X_c \quad \text{if } \sigma_1 < 0 \quad (29)$$

Inter-fiber failure (f_m) can be calculated based on different conditions differing three modes (Mode A, Mode B and Mode C). Matrix tension dominated (A) mode happens when the unidirectional stress perpendicular to the fiber direction is tension or non-existent ($\sigma_2 \geq 0$):

$$f_A = \sqrt{\left(\frac{\tau_{12}}{S_{12}}\right)^2 + \left(1 - p^+ \frac{Y_t}{S_{12}}\right)^2 \left(\frac{\sigma_2}{Y_t}\right)^2} + p^+ \frac{\sigma_2}{S_{12}} + (1 - f_w) \quad \text{if } \sigma_2 \geq 0 \quad (30)$$

$$p^+ = -\frac{d\tau_{12}}{d\sigma_2} \Big|_{\sigma_2=0} ; \sigma_2 \geq 0 \quad (31)$$

where p^+ is the slope of the failure envelope in $\tau_{12} - S_2$ plane, considered from $S_2 > 0$ stress quadrant.

Matrix shear dominant (B) mode is the following:

$$f_B = \frac{1}{S_{12}} \left(\sqrt{\tau_{12}^2 + (p^- \sigma_2)^2} + p^- \sigma_2 \right) + (1 - f_w) \quad (32)$$

$$\text{if } \sigma_2 < 0 \quad \text{and} \quad 0 \leq \left| \frac{\sigma_2}{\tau_{12}} \right| \leq \frac{R^A}{|\tau_{12,c}|}$$

$$p^- = - \left. \frac{d\tau_{12}}{d\sigma_2} \right|_{\sigma_2=0} ; \sigma_2 \leq 0 \quad (33)$$

$$R^A = \frac{S_{12}}{2p^-} \left(\sqrt{1 + 2p^- \frac{Y_c}{S_{12}}} - 1 \right) \quad (34)$$

$$\hat{p}^- = p^- \frac{R_A}{S_{12}} \quad (35)$$

$$\tau_{12,c} = S_{12} \sqrt{1 + 2\hat{p}^-} \quad (36)$$

where p^- is a slope similar to p^+ but considered from the $S_1 < 0$ quadrant, $\hat{p}^-$ stands for transverse-transverse slope, R^A is the transverse-transverse shear fracture resistance $\tau_{12,c}$ turning point shear stress.

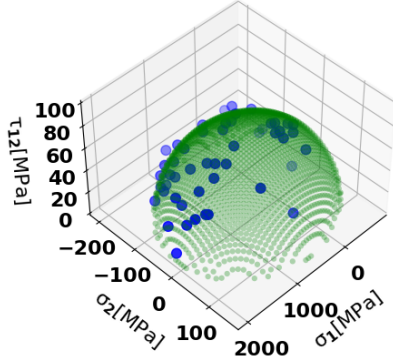
Matrix compression dominant failure (Mode C) when the shear stress is not significant:

$$f_C = \left[\left(\frac{\tau_{12}}{2(1 + \hat{p}^-)S_{12}} \right)^2 + \left(\frac{\sigma_2}{Y_c} \right)^2 \right] \frac{Y_c}{-\sigma_2} + (1 - f_w) \quad \text{if } \sigma_2 < 0 \quad \text{and} \quad 0 \leq \left| \frac{\tau_{12}}{\sigma_2} \right| \leq \frac{|\tau_{12,c}|}{R^A} \quad (37)$$

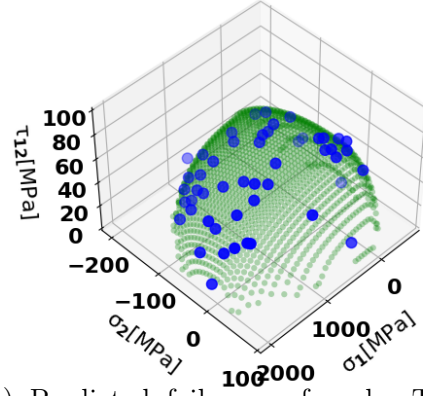
A failure occurs when the maximum of f_{ff} and f_m reaches value one.

$$f = \max(f_{ff}, f_m) \quad (38)$$

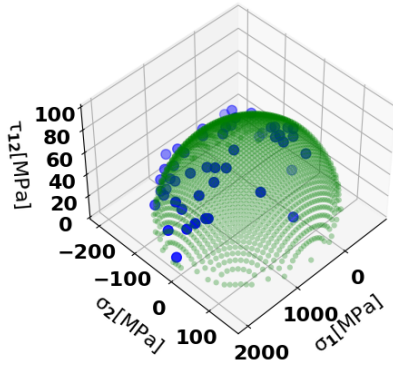
A.2 Failure surfaces of the use case



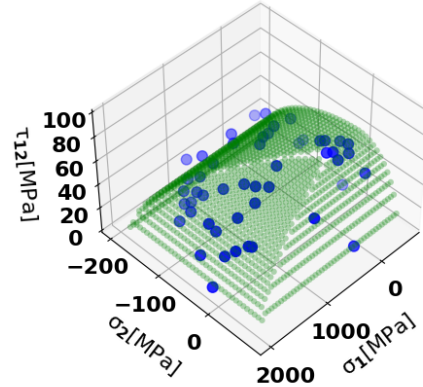
(a) Predicted failure surface by Tsai-Wu model



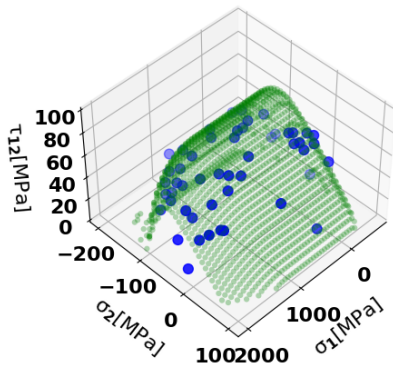
(b) Predicted failure surface by Tsai-Hill UD model



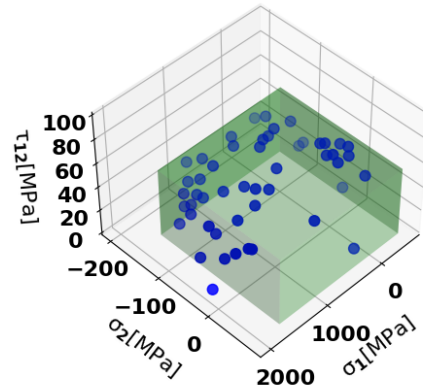
(c) Predicted failure surface by Hoffman model



(d) Predicted failure surface by Hashin model



(e) Predicted failure surface by Puck model



(f) Predicted failure surface by Max-Stress-Strain model

Figure 8: Predicted failure surfaces by Tsai-Wu, Tsai-Hill UD, Hoffman, Hashin, Puck and Max-Stress-Strain models in the model selection phase

References

- [1] Patricia L. Stumpff. Introduction to failure analysis. *Composites*, ASM Handbook, pages 949–950. ASM International, Materials Park, 2001.
- [2] M Ozyildiz, C Muyan, and D Coker. Strength Analysis of a Composite Turbine Blade Using Puck Failure Criteria. *Journal of Physics: Conference Series*, 1037:042027, jun 2018.
- [3] Libin Zhao, Wen Yang, Tiancheng Cao, Haibo Li, Baorui Liu, Chao Zhang, and Jianyu Zhang. A progressive failure analysis of all-C/SiC composite multi-bolt joints. *Composite Structures*, 202:1059–1068, 2018.
- [4] Mohammad S. Rouhi, Dimitra Ramantani, and Tong Earn Tay. 3D explicit simulation of bearing failure in metal–composite bolted joints. *Composite Structures*, 284:115108, 2022.
- [5] Miran Merhar. Application of Failure Criteria on Plywood under Bending. *Polymers*, 13(24), 2021.
- [6] C. Prakash and K. S. Vijay Sekar. Influence of Friction Coefficient and Failure Model in 3D FEA Simulation of Drilling of Glass Fiber Reinforced Polymer Composites. In *Advances in Manufacturing Processes*, pages 81–90, 2019.
- [7] Shuguang Li. The Maximum Stress Failure Criterion and the Maximum Strain Failure Criterion: Their Unification and Rationalization. *Journal of Composites Science*, 4(4), 2020.
- [8] Stephen W. Tsai and Edward M. Wu. A General Theory of Strength for Anisotropic Materials. *Journal of Composite Materials*, 5(1):58–80, 1971.
- [9] Oscar Hoffman. The Brittle Strength of Orthotropic Materials. *Journal of Composite Materials*, 1(2):200–206, 1967.
- [10] Stephen W. Tsai. Strength Theories of Filamentary Structures Fundamental Aspects of Fiber Reinforced Plastic Composites. *Wiley-Interscience*, 1968.
- [11] Z. Hashin. Failure Criteria for Unidirectional Fiber Composites. *Journal of Applied Mechanics*, 47(2):329–334, 06 1980.
- [12] A. Puck and H. Schürmann. Failure analysis of FRP laminates by means of physically based phenomenological models. *Composites Science and Technology*, 62(12):1633–1662, 2002.

- [13] Bhuiyan F.H., Kotthoff L., and III Fertig R.S. A machine learning technique to predict biaxial failure envelope of unidirectional composite lamina. 3:1451—1463, 2018.
- [14] Can B. Kalayci, Sevcan Karagoz, and Oezler Karakas. Soft computing methods for fatigue life estimation: A review of the current state and future trends. *Fatigue & Fracture of Engineering Materials & Structures*, 43(12):2763–2785, 2020.
- [15] Ghulam Mustafa, Afzal Suleman, and Curran Crawford. Probabilistic first ply failure prediction of composite laminates using a multi-scale M-SaF and Bayesian inference approach. *Journal of Composite Materials*, 52(2):169–195, 2018.
- [16] Doğa Gürgünöğlu, Berkan Dulek, and Sinan Gezici. Power adaptation for vector parameter estimation according to fisher information based optimality criteria. *Signal Processing*, 192:108–390, 2022.
- [17] Olivier Allix, Pierre Ladeveze, and E Vittecoq. Modelling and identification of the mechanical behaviour of composite laminates in compression. *Composites science and technology*, 51(1):35–42, 1994.
- [18] Joseph HA Guillaume, John D Jakeman, Stefano Marsili-Libelli, Michael Asher, Philip Brunner, Barry Croke, Mary C Hill, Anthony J Jakeman, Karel J Keesman, Saman Razavi, et al. Introductory overview of identifiability analysis: A guide to evaluating whether you have the right type of data for your modeling purpose. *Environmental Modelling & Software*, 119:418–432, 2019.
- [19] Peigen Yu, Mei Yin Low, and Weibiao Zhou. Design of experiments and regression modelling in food flavour and sensory analysis: A review. *Trends in Food Science Technology*, 71:202–215, 2018.
- [20] Gaia Franceschini and Sandro Macchietto. Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science*, 63(19):4846–4872, 2008.
- [21] James C Spall. Monte Carlo computation of the fisher information matrix in non-standard settings. *Journal of Computational and Graphical Statistics*, 14(4):889–909, 2005.
- [22] Dona Chatterjee, Arghya Ghosh, and Dipankar Chakravorty. Nonlinear first ply failure study of laminated composite skew plates. *Materials Today: Proceedings*, 45:4925–4930, 2021.

- [23] Dazhi Jiang and Dongwei Shu. Effects of damage thresholds on stresses in composite laminates under transverse impact. *Composite Structures*, 66(1):61–67, 2004.
- [24] Didier Dubois, Hélène Fargier, Meïssa Ababou, and Dominique Guyonnet. A fuzzy constraint-based approach to data reconciliation in material flow analysis. *International Journal of General Systems*, 43(8):787–809, 2014.
- [25] Ivan Markovsky and Sabine Van Huffel. Overview of total least-squares methods. *Signal Processing*, 87(10):2283–2302, 2007. Special Section: Total Least Squares and Errors-in-Variables Modeling.
- [26] Bussieck, Michael R and Pruessner, Armin *et al.* Mixed-integer nonlinear programming SIAG/OPT Newsletter: Views & News 14(1):19–22, 2003
- [27] Boxin Tang. Orthogonal array-based latin hypercubes. *Journal of the American Statistical Association*, 88(424):1392–1397, 1993.
- [28] Wes McKinney et al. Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference*, volume 445, pages 51–56, 2010.
- [29] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.
- [30] Matthew Newville, Till Stensitzki, Daniel B. Allen, and Antonino Ingargiola. LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python, 2014.
- [31] Per A. Brodtkorb. numdifftools package, 2021. URL: <https://pypi.org/project/numdifftools/>
- [32] Laurent Perron and Vincent Furnon. OR-Tools v9.4., 2022. URL: <https://developers.google.com/optimization/>.
- [33] Som R. Soni. A comparative study of failure envelopes in composite laminates. *Journal of Reinforced Plastics and Composites*, 2(1):34–42, 1983.