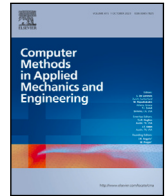


Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

## Comput. Methods Appl. Mech. Engrg.

journal homepage: [www.elsevier.com/locate/cma](http://www.elsevier.com/locate/cma)

# Incorporating sufficient physical information into artificial neural networks: A guaranteed improvement via physics-based Rao-Blackwellization

Gian-Luca Geuken, Jörn Mosler, Patrick Kurzeja \*

Institute of Mechanics, Department of Mechanical Engineering, TU Dortmund University, Leonhard-Euler-Str. 5, 44227 Dortmund, Germany

## ARTICLE INFO

## Keywords:

Artificial neural networks  
 Rao-Blackwell  
 Sufficiency  
 Information  
 Material modeling

## ABSTRACT

The concept of Rao-Blackwellization is employed to improve predictions of artificial neural networks by physical information. The error norm and the proof of improvement are transferred from the original statistical concept to a deterministic one, using sufficient information on physics-based conditions. The proposed strategy is applied to material modeling and illustrated by examples of the identification of a yield function, the identification of driving forces for quasi-brittle damage and rubber experiments. Sufficient physical information is employed, e.g., in the form of invariants, parameters of a minimization problem, isotropy and differentiability. It is proven how intuitive accretion of information can yield improvement if it is physically sufficient, but also how insufficient or superfluous information can cause impairment. Opportunities for the improvement of artificial neural networks are explored in terms of the training data set, the networks' structure and output filters. Even crude initial predictions are remarkably improved by reducing noise, overfitting and data requirements.

## 1. Introduction

## 1.1. Motivation: Connecting artificial neural networks and physical information

The aim of this work is to incorporate sufficient physical information into an artificial neural network (ANN) by a statistical concept, namely, Rao-Blackwellization. It stands out by the possibility to provide guaranteed improvements even for crude estimators. This statistical concept of Rao-Blackwellization remains to be explored from the viewpoint of deterministic physical modeling. In general, the combination of physical knowledge and analytical or numerical methods has nevertheless a rich tradition. The development of specialized finite elements, as one classical example, has often been driven by knowledge about the physical system of interest, e.g., reduced integration to avoid locking problems for nearly incompressible materials or shell structures with plane-stress states [1,2]. Also multi-scale problems can benefit from efficient information transfer, e.g., see a physics-based adoption of the Rao-Blackwell concept to capillary effects in [3]. The present focus of application is oriented towards ANN-based material modeling, because it benefits strongly from information-based improvement strategies.

More recently, the emerging field of data-based models and ANNs in particular are found to be enhanced by physical information. One of the first combinations of artificial neural networks and physical modeling is the work of Forsell and Lindskog [4] for tank level modeling. An early work in the field of constitutive modeling of solid mechanical systems is from Liang and Chandrashekhara [5] for elastomeric foams. Bhadeshia [6] stated that “it is not clear how a neural network trained on the output of physical models

\* Corresponding author.

E-mail address: [patrick.kurzeja@tu-dortmund.de](mailto:patrick.kurzeja@tu-dortmund.de) (P. Kurzeja).<https://doi.org/10.1016/j.cma.2024.116848>

Received 1 December 2023; Received in revised form 14 February 2024; Accepted 14 February 2024

Available online 19 February 2024

0045-7825/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

actually captures the principles of the physical models and more tests need to be done to study the long-range extrapolation behavior of such models”, indicating the rising interest in hybrid neural network modeling. The introduction of physics-informed neural networks (PINNs) [7,8] was a significant step to solve forward and inverse problems by involving nonlinear partial differential equations. Since then, several variations and extensions have been proposed such as the physics-informed PointNet (PIN) [9], extended physics-informed neural networks (XPINNs) [10] or the application to continuum micromechanics [11]. [12] provides a review on applications and variations of PINNs.

Linka et al. [13] also support data-driven constitutive modeling by continuum-mechanics knowledge in a unified network approach referred to as constitutive artificial neural networks (CANNs). Information on the physical problem helps to reduce the required data set as was also demonstrated by Tang et al. [14]. The data-driven approach suggested by Kirchdoerfer and Ortiz [15] and the application to parameter identification in [16] show a stronger orientation towards the material behavior, i.e., directly to experimental material data and pertinent constraints with conservation laws. A clear motivation for improvements of data-driven approaches can generally be seen in the traditional ideas of model order reduction, e.g., by training of scale separation [17] or reduced homogenization, e.g., for transient diffusion problems [18]. Settgest et al. [19] demonstrate the usefulness of a hybrid multi-scale neural network approach for the description of yield functions and evolution equations. This also applies to uncertain data and reliability analysis, for instance, during the design stage of artificial neural networks for real-time predictions of mechanized tunneling processes [20].

### 1.2. The more physical information, the better the prediction?

While utilizing such physical information can yield an improvement, this is not granted, though. Even worse, an unfortunate implementation of physically sound constraints into an ANN or other frameworks may lead to poorer predictions. Before a strategy is derived on the basis of Rao-Blackwellization, this should be introduced clearer by a simplified example. The final example of this study will later also discuss the now evident limitations for the practical use of digital image correlation (DIC) data with ANNs in numerical simulations.

Let us assume a nearly-incompressible polymer with Poisson’s ratio  $\nu$  below but close to 0.5. Experimental data of a stretched bicycle tube patch is shown in Fig. 1. Longitudinal strain ( $\epsilon_{xx}$ ) is determined from reliable displacement measurements at the clamped ends. The limited image resolution and edge detection, however, causes low sensitivity for lateral strain. Lateral strains ( $\epsilon_{yy}$ ) are thus obtained via DIC at multiple regions of the image. As a result, the  $\epsilon_{yy}$  component thus contains multiple yet noisy points for each stretched state, e.g., approximately of the format  $\epsilon_{yy} = -\nu\epsilon_{xx} + \epsilon$  with a normal-distribution error  $\epsilon \sim \mathcal{N}(0, \sigma^2)$ . Some of this data may even exceed the physically sound limits of incompressible and auxetic materials. In a first attempt, only the physically sound data base shall be treated as remaining sufficient information. This means disregarding data that exceeds the incompressible and auxetic limits as a pre-processing step ( $\nu = 0.5$  and  $\nu = 0$ , shaded regions in Fig. 1). Clearly, there exist established (statistical) methods for dealing with noisy data. Nevertheless, the straightforward implementation of a physically sound constraint should illustrate how it can impair the later prediction; by biasing the noisy data set in this introductory example.

Removing this unphysical data worsens the prediction for a subsequent linear regression. The Poisson’s ratio is  $\nu = 0.23$  for the reduced set of physically admissible data points. Accounting for all data, also the unphysical one, yields a more realistic prediction of  $\nu = 0.43$ . This is closer to the typical value of soft polymers. Obviously, simply imposing physically sound side constraints does not guarantee improvement. The accessibility of this introductory example fortunately allows to understand the obvious misuse of physical information. Applying the physical constraints to the noise  $\epsilon$  of the original data adds a bias, which affects later ANN training or fitting.

More complex physical systems described by ANNs, however, often do not allow such a clear interpretation. Sources of uncertainties are plentiful and so are the possibilities to implement physical information into ANNs, e.g., into their training data, their structure or their hyperparameters. To apply physical knowledge may be well justified. But asking

- what information to incorporate?
- when to incorporate this information?
- how to incorporate this information?

are questions that are not easily assessed. They motivate the present study. It will be shown how intuitive assumptions such as isotropy can help to improve the generation of ANNs, but also where limitations may arise if superfluous or biased information is transferred. The present approach specifically questions the unconditional benefits of physical information by providing a scheme that yields an improvement if sufficient information is incorporated in the form of a conditional average as illustrated in the later examples.

### 1.3. Rao-Blackwellization of physical models: sufficient information for guaranteed improvement

In addition to significant developments in physics-based ANNs, strategies of efficiently employing information have also been elaborated in the area of statistical modeling. The specific inspiration for the present work is the so-called Rao-Blackwell(-Kolmogorov) theorem [21–24]. It relies on knowledge about sufficient information. More precisely, it states that the conditional expectation of an initial estimator, given a sufficient statistic, is a better estimator. It can be proven mathematically that the new estimator, which employs sufficient information, is never worse. This optimality condition and the construction scheme for the new estimator will help to answer the questions raised above.

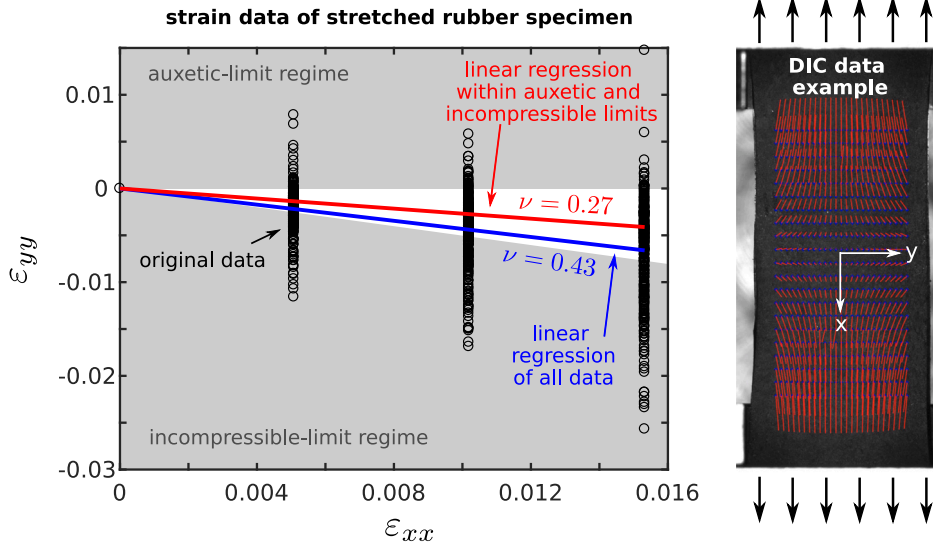


Fig. 1. Uniaxial stretch of a bike-wheel rubber sample: Poisson’s ratio approximation based on original data vs. data not exceeding incompressible and auxetic limits. An example of DIC-calculated displacements is given by red lines superposed on the sample on the right. Longitudinal strain ( $\epsilon_{xx}$ ) is determined from clamp displacement, lateral strain ( $\epsilon_{yy}$ ) is derived from DIC data over multiple regions.

It should be underlined that the Rao–Blackwell theorem is defined in a statistical framework and cannot be applied directly to a deterministic modeling framework. The present work hence combines the mathematical structure of Rao–Blackwellization with sufficient physics information for the enrichment of artificial neural networks. It is important to note that the proposed, new combination will harness information about the physical model [3] instead of statistical models [25]. Referring to the term Rao–Blackwellization should therefore not be misleading towards the original concept, but should acknowledge the original idea.

The focus will be on material models in the framework of solid continuum mechanics. By doing so, the presented approach complements existing approaches of physics-based ANNs in three different ways. It

- does not ask for more information but just for sufficient information,
- mathematically introduces a rigorous error measure and
- can be applied even to crude initial predictions.

A first result is clearly the sufficiency criterion, i.e., using just a sufficient set is preferable. Even more so, however, Rao–Blackwellization exceeds that intuitive guideline by providing a proof and a scheme for the improvement of estimators. This is considered particularly beneficial for scenarios with incomplete data sets and where the format of the solution is secondary, i.e., data-based solutions such as ANNs used in practice.

Before we present the structure of the following chapters, we find it helpful to distinguish the proposed approach from other statistical methods. The present approach combines three research areas: the problem of deterministic physical models, ANNs as a means for their prediction and Rao–Blackwellization for a guaranteed improvement. The adopted concept of Rao–Blackwellization emerges from frequentist inference, which matches the idea of deterministic physical models with fixed and unique parameters. The consideration of uncertain parameter distributions exceeds the present focus and is often accounted for by Bayesian approaches. Gaussian process regression models, for example, allow the implementation of Gaussian processes [26] in the form of a nonlinear, nonparametric regression tool [27] in the area of data-based material modeling [28,29]. Methods for model discovery from regression analysis such as symbolic regression [30,31] also share the goal of variable selection. In contrast to the current employment of external/physics-based knowledge, however, they are oriented towards extracting the underlying relationships from the data set. Lasso (least absolute shrinkage and selection operator) regularization [32,33] also aims at the discard of irrelevant or redundant variables by coefficient penalization. It does not follow the presented scheme of physics-based conditional averaging, though, but instead can support feature selection and interpretable models based on the available data. Within the statistical context of the original Rao–Blackwell theorem, the transfer of further properties like completeness and unbiased predictions may even lead to uniqueness and optimality of the improvement in the sense of the Lehmann–Scheffé theorem [22]. A more comprehensive transfer to physical problems may then be achieved for applications like homogenization theory when inferring macroscopic properties from the induced microscopic distributions; for instance, relating sufficiency to injective information transfer across the scales, completeness to surjectivity and unbiased predictions to convergence based on representative volume elements and scale separation laws. Another distinction shall be emphasized with respect to the general incorporation of physics-based constraints. The Rao–Blackwell scheme is intended to improve existing predictions using sufficient information. The scheme itself and its purpose hence complement but do not

substitute approaches that employ constraints for the initial generation of ANNs (e.g., penalization or special architectures [12,13]) or other types of physics-related information (e.g., the second law of thermodynamics).

The present work will first establish a formal framework for Rao-Blackwellization of physical models in Section 2. It includes an introductory example based on the micro-sphere model and introduces general concepts for ANNs. The general concepts are then further investigated and highlighted by three examples. Section 3 provides an analytical example for predicting a yield function utilizing a description via invariants. It points out how physics-based Rao-Blackwellization can cure noisy data as well as under- or overfitted ANN models, if applied to the ANN output. A damage model based on a variational minimization problem allows to compare different realizations of the proposed improvement strategy in Section 4 in terms of modifications of ANN output, data and structure. Finally, the numerical adoption in Section 5 utilizes isotropic elasticity for the simulation of rubber samples based on DIC and force–displacement data. This example emphasizes data processing and augmentation via physics-based Rao-Blackwellization. Overall, each example highlights different aspects of how sufficient information can improve ANNs from data generation, to the ANN structure and its output.

## 2. Transferring Rao-Blackwellization from statistical models to deterministic material models

### 2.1. Adaption of the Rao-Blackwellization framework to material models

Statistical modeling and physical material modeling both share the main challenge of predicting unknown quantities. They are also similar as they both follow the strategy of employing a governing set of mathematical equations supported by modeling assumptions and assessed by error norms. The specific implementations and nomenclature yet differ, which is why we first explain the key differences that are relevant to adapt the Rao–Blackwell theorem.

While a statistical model builds upon a sample space (e.g. tensile specimen copies), the material model builds upon physical, deterministic configurations (e.g. stress states or microstructures). Statistical model parameters can be mean and standard deviation of sustained force, whereas material model parameters can be Young’s modulus and Poisson’s ratio. In the present context, these are regarded to be fixed deterministic values. On the one hand, a statistical estimator predicts an estimate of a given quantity from observed data (e.g. a sample mean force). On the other hand, an analogy in deterministic material modeling may be the prediction of an unknown Poisson’s ratio but also predictions of stresses or strains work within this framework. Finally, a possible error measure of a statistical estimator can be defined by the expected value  $\mathbb{E}$  of the mean squared error. A similar mean squared error can be introduced in material modeling by performing an average  $\mathbb{A}$  of the mean squared error over a set of physical configurations. The specific, necessary mathematical definitions and implementation are derived in the following and the interested reader is referred to [3] for more details, where physics-based Rao-Blackwellization is applied to information transfer in multi-scale problems instead of ANN-based processing.

To become more specific in terms of material modeling, we start from a prediction of an unknown quantity  $\theta$  that depends on the physical state  $\omega \in \Omega$ . For instance, we may want to distinguish elastic loading ( $\theta < 0$ ) from plastic loading ( $\theta = 0$ ) and from inadmissible stress states ( $\theta > 0$ ) for a given stress tensor ( $\omega = \sigma$ ). Moreover, we assume from a physical constraint that there is sufficient information  $S$  for the prediction of the unknown. For instance, we can assume that the effective macroscopic plastic yielding of polycrystals can often be approximated on the basis of the deviatoric stress norm ( $S = \|\sigma^{\text{dev}}\|$ ). Note that we do not need to know how the exact relationship  $\theta(\omega)$  looks like. The aim is to find a better one, which is achieved in three steps:

1. Let  $\theta_0(\omega)$  be any starting prediction.
2. Let  $S$  be sufficient information to determine  $\theta$ .
3. Then, the new prediction

$$\theta_1(s) = \mathbb{A}(\theta_0 | S = s) = \frac{1}{|\Omega_s|} \int_{\Omega_s} \theta_0(\omega) \, d\omega, \quad (1)$$

is the conditional average value of  $\theta_0$ ’s predictions for all  $\omega$  such that  $S(\omega) = s$ . This prediction is never worse in the sense of the averaged mean squared error

$$\frac{1}{|\Omega|} \int_{\Omega} (\theta_1 - \theta)^2 \, d\omega \leq \frac{1}{|\Omega|} \int_{\Omega} (\theta_0 - \theta)^2 \, d\omega. \quad (2)$$

Therein,  $\Omega_s := \{\omega \in \Omega | S(\omega) = s\}$  is the subdomain, where  $S(\omega)$  takes the constant value  $s$ .  $|\Omega| := \int_{\Omega} 1 \, d\omega$  and  $|\Omega_s| := \int_{\Omega_s} 1 \, d\omega$  are used for normalization.

In brief,  $\theta_1$  may be interpreted as a conditional averaging of the initial prediction  $\theta_0$  over constant values  $s$  of the sufficient information  $S$ . This interpretation is different from the statistical use of an expected value. What is most important, the physical constraint tells us that the prediction of  $\theta$  must be fully determined by prescribing  $s$ . For instance, two stress states  $\sigma_I$  and  $\sigma_{II}$  must be both elastic or both plastic if they have the same deviatoric stress norm,  $\|\sigma_I^{\text{dev}}\| = s_I = s_{II} = \|\sigma_{II}^{\text{dev}}\|$ . The adopted Rao–Blackwell theorem thus averages over such states for the new prediction. As a first key result, we did not only obtain a strategy to incorporate physical knowledge — it is guaranteed not to worsen our initial prediction. A fundamental property of this scheme is that the error does not grow. The proof is sketched in [Appendix A](#).

## 2.2. Notes on the practical implementation of sufficient information

The strategy employs three steps – initial estimator, sufficient set, improved estimator – each with individual practical peculiarities. Various initial predictions can be tested in application, e.g., neural networks trained with different initial weights, data or transfer functions. The initial prediction can even be a crude one, while two special cases would trivially not yield an improvement in the sense of a strict inequality (2); namely, the correct prediction (it cannot be improved further) and an already improved prediction (due to idempotency).

For the second step of the scheme, identifying a sufficient set is just the necessary information for the improvement scheme to properly work. While sufficiency in statistics is proven based on the statistical model properties (c.f. the Fisher-Neyman factorization [24,34]), sufficiency for the material model's predictions is based on knowledge about the physical behavior. Note that sufficiency does not imply sufficient information in every detail. Instead and like the statistical motivation, sufficiency is understood with respect to a model. Obvious sources determining a sufficient set in material modeling can be assumptions like isotropy, linearity or hyperelasticity. Conservation laws, observer invariance or the second law of thermodynamics can also reveal whether a parameter is allowed to have an impact on another. Other sources of sufficiency can be the parameters of a set of governing equations, which sufficiently fix the unknown solution. A dimensional or homogenization analysis can moreover help to reduce the set of required parameters for prediction. This list is not exhaustive and shall provide a few ideas for practical application.

Having sufficient information at hand, Eq. (1) defines the third step, i.e., the improved prediction. In the ideal case, the integral in (1) is evaluated analytically but may be unattainable in practice. In such cases, a numerical approximation of the scheme is required. The idea is to transform the continuous parameterization of  $s$  into discrete intervals. The discretization highly depends on the range of  $s$ , which may be unbounded or bounded by physical constraints. An intuitive approach for the latter case is to introduce  $n + 1$  intervals

$$s_i = \left[ s_{\min} + \frac{i}{n+1} (s_{\max} - s_{\min}), s_{\min} + \frac{i+1}{n+1} (s_{\max} - s_{\min}) \right), \quad (3)$$

with  $i = 0, \dots, n$ . With these at hand, the integral can be replaced by a sum over all data points  $m$ , which fall into the interval  $s_i$ , to determine the new model

$$\theta_1(s_i) = \frac{1}{m} \sum_{j=1}^m \theta_0(S(\omega) \in s_i). \quad (4)$$

As one result, the numerical approximation of the improved prediction may not fully inherit the optimality condition. It remains a compromise between interval stepping and averaged data, limited by data availability and computational capacities.

## 2.3. Introductory example for Rao-Blackwellization of physical constraints: the micro-sphere model

An example from the area of micro-sphere modeling [35] of rubber shall make the concept of Rao-Blackwellization with physical constraints comprehensible. It is purely analytical and not combined with artificial neural networks or numerical evaluations yet. Statistical considerations of micro-sphere modeling (such as Langevin statistics and random-walk of chain segments) are also omitted for better distinction of physics-based Rao-Blackwellization from statistics-based Rao-Blackwellization. Like all examples presented herein, the introductory example follows the same three steps:

1. initial estimator  $\theta_0$ ,
2. sufficient information  $S$ ,
3. improved estimator  $\theta_1$ .

For the sake of simplicity, the symbols  $\theta_0$ ,  $S$  and  $\theta_1$  will be used in all examples without further modification of the notation.

Consider a rubber network of polymer chains that undergoes macroscopic deformation described by the right Cauchy–Green tensor  $C$ , Fig. 2. The task is to predict an effective, quadratic stretch of the underlying polymer chains  $\lambda_c^2$ . This effective stretch is commonly required for the energy functions of rubber models [35,36]. We now have to pick some information from  $C$  to predict  $\lambda_c^2$ .

While there are a few obvious ad-hoc choices, some even considering superposed microstructural fluctuations, we deliberately start with a poor initial estimator. The initial estimator predicts the effective quadratic stretch of the polymer chains by the 11-component only of some Cartesian coordinate system

$$\theta_0 = C_{11} = \mathbf{e}_1 \cdot C \cdot \mathbf{e}_1. \quad (5)$$

This initial estimator is the first step of Rao-Blackwellization but only acceptable for uni-axial stretch. It clearly fails for biaxial tension, for instance. We would therefore like to reduce the error for all possible deformation states.

We now further assume that the assumption of isotropy is a valid physical approximation, e.g., based on fabrication conditions and experimental indication. The second step of Rao-Blackwellization then requires the sufficient information under this constraint. For an isotropic model, it is sufficient to know the three (rotational) invariants of the strain tensor  $C$ . These invariants thus constitute the sufficient set

$$S = \{I_1^C, I_2^C, I_3^C\}. \quad (6)$$

### Rao-Blackwellization of effective chain stretch

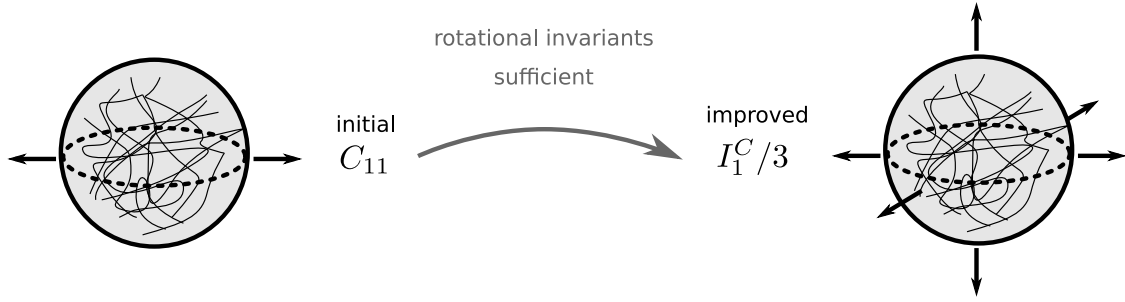


Fig. 2. Example of Rao-Blackwellization within the chain-based micro-sphere framework: chain stretch from one in three directions. Improvement is based on sufficient information under the assumption of rotational invariance.

The third step of Rao-Blackwellization now yields the improved prediction by averaging over the domain of deformation states, where the sufficient set is constant. This subdomain of deformation states, where  $S = \{I_1^C, I_2^C, I_3^C\}$  is constant, is given by just the rotations  $\mathbf{Q}$  of the strain tensor:

$$\begin{aligned} \Omega_s &:= \{\text{all } \mathbf{C} \text{ with same values of invariants } I_1^C, I_2^C, I_3^C\} \\ &= \{\mathbf{Q}^T \cdot \mathbf{C} \cdot \mathbf{Q} : \forall \mathbf{Q} \in \text{SO}(3) \text{ and starting } \mathbf{C} \text{ with invariants } I_1^C, I_2^C, I_3^C\}. \end{aligned}$$

The improved estimator hence reads (see [35] for the algebra of the last step)

$$\begin{aligned} \theta_1(I_1^C, I_2^C, I_3^C) &= \frac{\int_{\Omega_s(I_1^C, I_2^C, I_3^C)} \theta_0(C_{11}) d\omega}{\int_{\Omega_s(I_1^C, I_2^C, I_3^C)} 1 d\omega} = \frac{\int_{\Omega_s(I_1^C, I_2^C, I_3^C)} \mathbf{e}_1 \cdot \mathbf{C} \cdot \mathbf{e}_1 d\omega}{\int_{\Omega_s(I_1^C, I_2^C, I_3^C)} 1 d\omega} \\ &= \frac{\int_0^{2\pi} \int_0^\pi \mathbf{e}_1 \cdot \mathbf{Q}^T(\vartheta, \varphi) \cdot \mathbf{C} \cdot \mathbf{Q}(\vartheta, \varphi) \cdot \mathbf{e}_1 \sin \vartheta d\vartheta d\varphi}{\int_0^{2\pi} \int_0^\pi 1 \sin \vartheta d\vartheta d\varphi} \\ &= \frac{\int_0^{2\pi} \int_0^\pi (\mathbf{Q}(\vartheta, \varphi) \cdot \mathbf{e}_1) \cdot \mathbf{C} \cdot (\mathbf{Q}(\vartheta, \varphi) \cdot \mathbf{e}_1) \sin \vartheta d\vartheta d\varphi}{4\pi} \\ &= \frac{C_{11} + C_{22} + C_{33}}{3} = \frac{I_1^C}{3}. \end{aligned} \quad (7)$$

The improved estimator thus simplifies to a third of the first invariant  $\theta_1 = I_1^C/3$ . This is just the wanted effective, quadratic stretch  $\lambda_c^2$  for the limit case of the eight-chain Arruda-Boyce model when derived from the micro-sphere framework [35,36]. Application to other powers of the stretch, e.g. a linear estimation, may yield to different weighting factors and we refer to [35] for details on the micro-sphere approach. Moreover, [37] contains an alternative analytical example for rubber elasticity based on the formulation of principal stretches and [38] for the contraction of nanowires.

Rao-Blackwellization however stands out from a micro-sphere approach by some significant characteristics and another scope of application. Instead of developing an entire framework, it just employs an initial estimator (even an improper one is possible) and knowledge about sufficient information on the unknowns (which is the point where the additional physics enter the prediction). This makes it a suitable tool for situations where a complete physical framework is not provided, e.g., predictions based on artificial neural networks as demonstrated by the following scenarios.

#### 2.4. Incorporating sufficient physical information into ANNs via Rao-Blackwellization

The proposed physics-based Rao-Blackwellization implies the use of conditional averages based on a sufficient set of information, but it can be applied to an ANN framework in various ways. Three characteristic possibilities are chosen to illustrate this process. All of them will be demonstrated in detail in the following examples, but the main ideas shall already be sketched here, too:

- Filtering the ANN output:

The first option, filtering the ANN output, is directly in line with the three steps of Rao-Blackwellization. The initial prediction  $\theta_0$  is the ANN output given the physical state  $\omega$  as input. Knowing sufficient information  $S$  for the problem at hand, Eq. (1) defines the improved prediction  $\theta_1$ . Since the analytical integration of ANN outputs can become impractical, a numerical approximation is often necessary as described in Section 2.2. This strategy is an option for scenarios, in which neither the ANN nor the data can be modified.

- ANN training data augmentation:

Rao-Blackwellization also allows to directly work on the training, validation and test data of an ANN as a second option. It implies that the initial prediction of the ANN should yield the same output for data points that share the same values

of sufficient parameters. For instance, to be more precise, the Rao–Blackwell theorem offers training data augmentation to improve a given data set if the number of parameters  $n_S$  of  $S$  is smaller than the number of parameters  $n_\omega$  of  $\omega$ . Less information would be needed. In that case  $n_\omega - n_S$  parameters of  $\omega$  can be chosen freely for a fixed value  $S(\omega) = s$ . This allows to augment additional data points in the  $\Omega$  space. They will differ in terms of the input data format, but they will share the same sufficient information and hence the same output labels as the original data points they have been derived from. This option is applicable and not costly if the ANN structure is fixed, e.g., in existing numerical frameworks, but when the data can be modified.

- Adapting the ANN structure:

The third option modifies the initial ANN structure and replaces the input layer of the ANN to directly process the set of sufficient information  $S$ . This automatically ensures a constant output for the sufficient set  $S$ , because fluctuations are not possible by design. It is referred to as adapting or optimizing the ANN structure in this study. This fundamental adaption may be regarded to be the strongest enforcement of the Rao-Blackwellization concept. It does not necessarily yield the best results, though. A counterexample is given in Section 4, where all three options are compared.

### 3. Illustrating Rao-Blackwellization effects on an ANN - Example of an elasto-plastic predictor step

To introduce the application of the proposed output filter strategy to an ANN, the yield function defining the transition between a fully elastic and an elasto-plastic state is to be identified. Accordingly, the ANN shall predict whether a deformation will fall into the elastic regime or not. Point of departure is the classic von Mises yield criterion in the two-dimensional stress space

$$\Phi^{\text{VM}}(\sigma_1, \sigma_2) = \sqrt{\sigma_1^2 + \sigma_2^2 - \sigma_1 \sigma_2} - \sigma_Y \quad (8)$$

that is formulated in the principal stresses  $\sigma_1$  and  $\sigma_2$  and can be motivated for shear-dominated plasticity in crystals. The analytical starting point will also serve as a reference solution and eases an interpretation of the later improvement. The material's yield stress is chosen to be  $\sigma_Y = 1$  MPa. Considering an elastic predictor step within a numerical implementation (e.g., the return-mapping scheme), the different states can be identified as

$$\Phi^{\text{VM}}(\sigma_1, \sigma_2) \rightarrow \begin{cases} \leq 0 & \text{(elastic)} \\ > 0 & \text{(inadmissible} \Rightarrow \text{elasto-plastic corrector)}. \end{cases} \quad (9)$$

#### 3.1. Initial ANNs distinguishing elasticity from plasticity

The initial ANN,  $\theta_0$ , shall predict whether a given stress pair  $(\sigma_1, \sigma_2)$  corresponds to a purely elastic step or not. We test three initial ANNs of different sizes, which are layer-wise: 2-5-1, 2-10-5-1 and 2-200-200-60-1. The aforementioned stress pair is the input and all neurons are fully connected to the neighboring layers and activated by the hyperbolic tangent (tanh) function. The output of the last layer is moreover processed such that the discrete distinction between elasticity and plasticity is numerically labeled by 0 and 1 in the form (cf. Eq. (9))

$$o = \begin{cases} 0 & \text{if } \theta_0(\sigma_1, \sigma_2) \leq 0 \quad \text{(elastic)} \\ 1 & \text{if } \theta_0(\sigma_1, \sigma_2) > 0 \quad \text{(inadmissible} \Rightarrow \text{elasto-plastic corrector)}. \end{cases} \quad (10)$$

Data was generated from the analytical solution in the range of  $\Omega = [-1.75 \text{ MPa}, +1.75 \text{ MPa}]^2$  for training (144 data points), validation (144 data points) and testing, see Fig. 3. Training and validation data was moreover superposed with random noise to highlight the influence of imperfect data. Stress states that fall into the range  $[-0.03 \text{ MPa}, +0.03 \text{ MPa}]$  of the yield function were assigned randomly. For the test data, the stress domain  $\Omega$  was discretized by  $\Delta\sigma = 0.01$  MPa, leading to 123 201 test data points. No regularization of the input data is required in this example by working with units of MPa. The networks have been finally trained via the Stochastic Gradient Descent (SGD) method for 1500 epochs. The network performance was evaluated by the mean squared error of the respective data sets during training and testing. A detailed summary of the ANN hyperparameters and settings is provided in Appendix B for all examples. The initial ANN of size 2-5-1 shows a significant error due to its simplicity that causes underfitting. Unphysical predictions can be observed in the form of separated and non-convex elastic regimes, see the initial (gray) performance in Fig. 4. The 2-10-5-1 network performs much better, matching both the qualitative and the quantitative prediction of the elastic–plastic regimes. The 2-200-200-60-1 network achieves an even lower error in the training set, but the validation and test set errors are higher compared to the 2-10-5-1 case. This overfitting behavior is caused by the high number of degrees of freedom. Given the initial prediction of all networks ( $\theta_0$ ), all three shall be optimized via physics-based Rao-Blackwellization.

#### 3.2. Sufficient physical information for the yield condition

Motivated by slip in crystals, for instance, we assume that the onset of plasticity is dictated by the deviatoric stress norm. For the present principal stresses and analytical reference, it writes

$$S(\sigma_1, \sigma_2) := \|\sigma^{\text{dev}}\| = \sqrt{\sigma_1^2 + \sigma_2^2 - \sigma_1 \sigma_2} \quad (11)$$

and completely determines whether the material undergoes elastic or plastic deformation. It is thus *sufficient* to distinguish between both cases under this assumption. The simple, underlying analytical solution in (9) is provided here as an idealized benchmark to be met for better interpretation of the results. Unpredictable solutions follow in the later examples. Also note at this point that we do not need to know how this sufficient information relates to the prediction. The Rao–Blackwell scheme only receives an imperfect ANN as an initial predictor and the sufficient set of information.

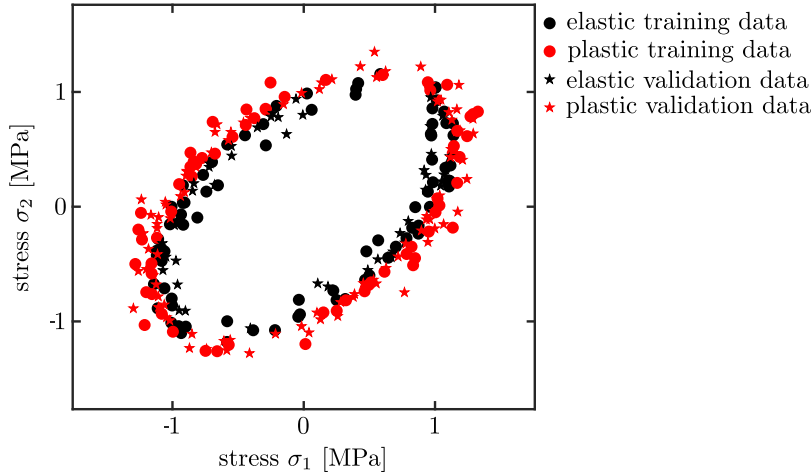


Fig. 3. Visualization of training and validation data for the ANNs distinguishing elastic and plastic predictor steps. The test data covers the entire depicted stress domain with a fine discretization of  $\Delta\sigma = 0.01$  MPa and is therefore not visualized by markers.

### 3.3. Improved ANNs distinguishing elasticity from plasticity

The improved ANNs,  $\theta_1$ , are calculated as, cf. (1),

$$\theta_1(\sigma^{\text{dev}}) = \frac{1}{|\Omega_{\sigma^{\text{dev}}}|} \int_{\Omega_{\sigma^{\text{dev}}}} \theta_0 \, d\sigma, \quad (12)$$

$$\Omega_{\sigma^{\text{dev}}} = \{ \{\sigma_1, \sigma_2\} \in [-1.5 \text{ MPa}, +1.5 \text{ MPa}]^2 \text{ with } \|\sigma^{\text{dev}}(\sigma_1, \sigma_2)\| = \sigma^{\text{dev}} \}. \quad (13)$$

They hence only depend on a single scalar value  $\sigma^{\text{dev}}$ . They derive from the initial ANNs (depending on  $(\sigma_1, \sigma_2)$ ) by conditional averaging over all stress states with the same effective stress value  $\sigma^{\text{dev}}$ , which are denoted by the subdomain  $\Omega_{\sigma^{\text{dev}}}$ . Potential reduction of the input parameters is already an important outcome for the application to neural networks.

The initial ANNs provide discrete distributed data but no solution that allows for an explicit analytical integration. We thus approximate the stress domain of the improvement scheme numerically by a discretization with intervals  $\sigma_i^{\text{dev}} = [(i - 0.5) \, ds, (i + 0.5) \, ds]$ ,  $i = 0, \dots, n$ . The integral is then replaced by a sum over all  $m$  data points that fall into the range of  $\sigma_i^{\text{dev}}$ . Interpolation is used for empty intervals and the interval number is limited to  $n = 1750$ . Just like for the initial ANNs, the output of the new ANNs is also rounded to 0 or 1 to obtain a clear distinction between elastic and plastic predictors, respectively.

Physics-based Rao-Blackwellization improves all initial networks' output (sizes 2-5-1, 2-10-5-1 and 2-200-200-60-1) significantly in terms of the cost function and the shape of the flow curves, Fig. 4. Underfitting of the smallest network initially caused completely unphysical predictions that are now well solved by a single, convexified elastic domain. The error dropped by a factor of 8. Also overfitting of noise in the largest network is now successfully suppressed by an error improvement of more than factor 3. Even the very good prediction of the middle-sized network was further improved notably by a factor of 15 in terms of the cost function. The physics-based Rao-Blackwell strategy hence shows clear improvements of all ANNs. Smoothing of noise and overfitting are two illustrative outcomes in this first example.

## 4. Improvement of ANNs by output filtering vs. data augmentation vs. structure adaption: example of quasi-brittle damage

Rao-Blackwellization can be applied at different steps of the ANN evolution, though practical operational conditions may limit these options. On the one hand, the ANN structure may be fixed when embedded into a standardized software package or prespecified by convexity or symmetry requirements [39,40]. On the other hand, further data acquisition may be unavailable due to confidentiality or experimental limitations. Even intuitive design approaches such as data pre-/post-processing may be restricted for various practical reasons. The presented physics-based Rao-Blackwellization scheme must then be adopted to the options of the problem at hand. We will compare three different options in this example as introduced in Section 2.4, namely, filtering the initial ANN output, augmentation of the training data and adaption of the ANN structure.

To compare the three different options for improvement, we employ ANNs to the modeling of quasi-brittle damage. A common formulation for the energy density of an elastic material undergoing quasi-brittle damage, e.g., for phase-field formulations, depends on strain  $\epsilon$  and damage variable  $\alpha$  as [41]

$$W = \psi(\epsilon, \alpha) + w_1 (w(\alpha) + l^2 \|\nabla \alpha\|^2). \quad (14)$$

The first term is the elastic energy density decreasing with damage  $\alpha$ . The terms inside the brackets correspond to the fracture energy as far as phase-field theory is concerned. In a more general setting, it is responsible for regularization, cf. [42]. Basic phase-field

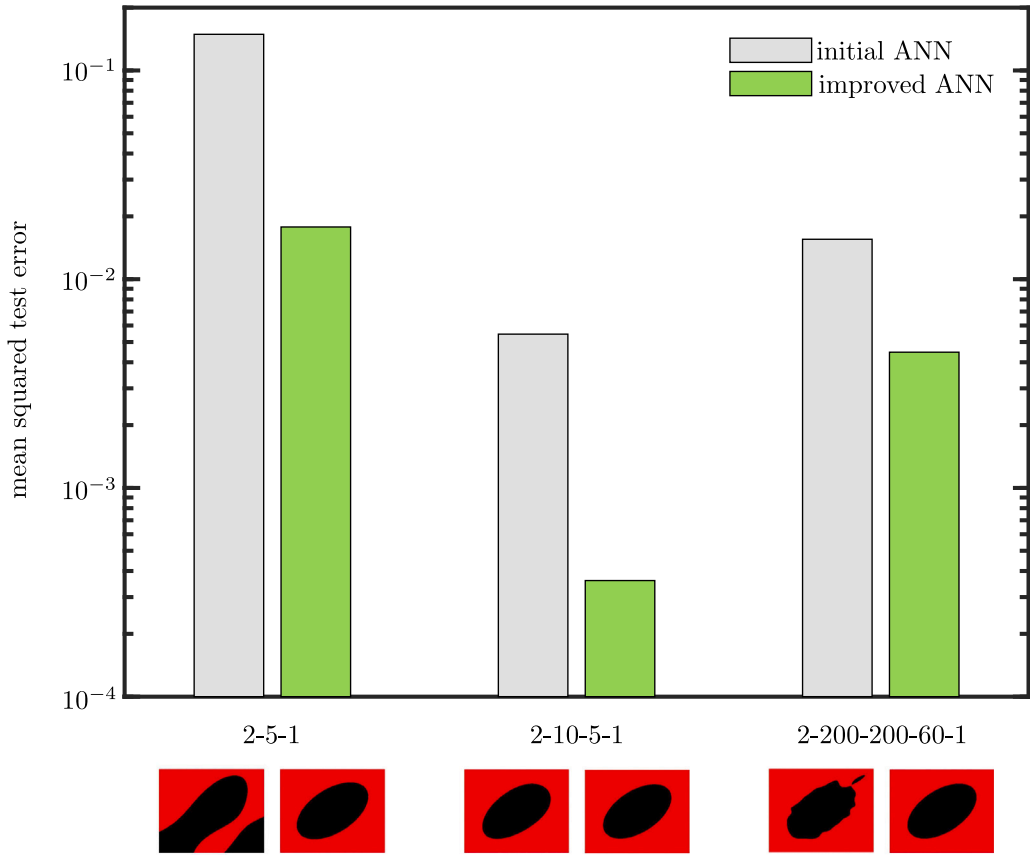


Fig. 4. Mean squared errors determined by the test set and yield functions for different network sizes before and after Rao-Blackwellization.

fracture models however still require an adaption to capture asymmetric tension-compression behavior. We follow the approach of De Lorenzis and Maurini [41] by a decomposition of the elastic energy that also determines the driving force for the damage variable. It takes the form

$$\psi(\boldsymbol{\varepsilon}, \alpha) = a(\alpha) \psi_D(\boldsymbol{\varepsilon}) + \psi_R(\boldsymbol{\varepsilon}) \quad \text{with} \quad \psi_D(\boldsymbol{\varepsilon}) + \psi_R(\boldsymbol{\varepsilon}) = \psi_0(\boldsymbol{\varepsilon}). \quad (15)$$

Therein,  $\psi_R$  is the residual elastic energy and only  $\psi_D$  is coupled to damage. In the theory of structured deformations [43], the elastic energy of the undamaged material  $\psi_0(\boldsymbol{\varepsilon})$  is decreased by the evolution of inelastic deformations  $\boldsymbol{\eta}$ , the so-called structured deformations. They are constrained to be in a convex set  $\mathcal{K}_\varepsilon$  in order to account for the structure of admissible micro-cracks. The elastic energy of the fully damaged body, given  $\psi_0(\boldsymbol{\varepsilon})$  and  $\mathcal{K}_\varepsilon$ , is found by the minimization problem:

$$\bar{\boldsymbol{\eta}}(\boldsymbol{\varepsilon}) = \arg \min_{\boldsymbol{\eta} \in \mathcal{K}_\varepsilon} \psi_0(\boldsymbol{\varepsilon} - \boldsymbol{\eta}), \quad (16)$$

$$\psi_R(\boldsymbol{\varepsilon}) = \min_{\boldsymbol{\eta} \in \mathcal{K}_\varepsilon} \psi_0(\boldsymbol{\varepsilon} - \boldsymbol{\eta}) = \psi_0(\boldsymbol{\varepsilon} - \bar{\boldsymbol{\eta}}(\boldsymbol{\varepsilon})). \quad (17)$$

Accordingly, the idea is similar to that of eigenfracture, cf. [44]. Under certain assumptions, such as a linear elastic and isotropic pristine material behavior, the minimization problem can be specified to

$$\bar{\boldsymbol{\eta}}(\boldsymbol{\varepsilon}) = \arg \min_{\boldsymbol{\eta}} \left\{ \frac{\kappa}{2} \text{tr}(\boldsymbol{\varepsilon} - \boldsymbol{\eta})^2 + \mu \|\boldsymbol{\varepsilon}_{\text{dev}} - \boldsymbol{\eta}_{\text{dev}}\|^2, \boldsymbol{\eta} \in \text{Sym}: \text{tr}(\boldsymbol{\eta}) \geq \gamma \|\boldsymbol{\eta}_{\text{dev}}\| \right\}, \quad (18)$$

wherein  $\kappa$  is the bulk modulus,  $\mu$  is the shear modulus and  $\gamma$  is a free parameter, which controls the ratio between the compressive and the tensile strength in uni-axial loading conditions. The ANNs are employed to predict the resulting energy proportions,  $\psi_D$  and  $\psi_R$ . The training data is obtained from the solution provided in [41] and, for better comparison, we assume plane strain-states from now on.

#### 4.1. Initial ANN structure for energy prediction

All of the following ANNs have a network size of 3-50-50-50-20-2, except for the structure-optimized ANNs (see Section 4.3), with a Rectified Linear Unit (ReLu) activation function and are trained for 10 000 epochs via the SGD method with a learning rate

of 0.01 and a batch size of 1, in order to ensure comparability. Also, five ANNs have been trained for each variant to account for random weight initialization and local minima during training.

The initial ANN takes the strain components  $\epsilon_{11}$ ,  $\epsilon_{22}$  and  $\epsilon_{12}$  as input parameters. The test data set comprises all possible combinations of the three strain components in the range of  $[-0.1, +0.1]$  with a discretization of  $\Delta\epsilon = 0.001$ . Due to the view on practical constraints, three different training data sets with a different amount of data points were generated based on the three strain components, with each component in the range of  $[-0.1, +0.1]$  for comparison. The data distribution is illustrated in Fig. 5 (bottom).

#### 4.2. Sufficient set from variational minimization problem

As an additional focus of investigation, this example will highlight the role of sufficiency. Three types of sufficient information will be distinguished for the variant of structure improvement. Firstly, non-sufficient information lacks some part of the sufficient set. Secondly, sufficient information fully determines the unknown parameter. Thirdly, minimally sufficient information is not only sufficient but cannot be reduced further without becoming non-sufficient.

From the minimization problem (18) one can conclude that one sufficient set is given by

$$S = \{\|\epsilon_{\text{dev}}\|, \text{tr}(\epsilon)\}. \quad (19)$$

Fixing the deviatoric norm and the trace of the strains will determine the evolution of the inelastic deformations and finally the individual energy contributions  $\psi_R$  and  $\psi_D$ . Again, we do not need to know how this relationship looks like. This moreover allows to specify a non-sufficient, a minimal sufficient and a non-minimal sufficient set, reading

$$S_1 = \{\|\epsilon_{\text{dev}}\|\} \quad (\text{non-sufficient}), \quad (20)$$

$$S_2 = \{\|\epsilon_{\text{dev}}\|, \text{tr}(\epsilon)\} \quad (\text{minimal sufficient}), \quad (21)$$

$$S_3 = \{\|\epsilon_{\text{dev}}\|, \text{tr}(\epsilon), \epsilon_{11}\} \quad (\text{non-minimal sufficient}). \quad (22)$$

Note that the initial ANN input is also a non-minimal sufficient set  $S_4 = \{\epsilon_{11}, \epsilon_{22}, \epsilon_{12}\}$ .

#### 4.3. Improved output, data and structure in comparison

##### 4.3.1. Methodological implementation

The first option for improvement by Rao-Blackwellization is the output filter. The ANN output is improved by a conditional averaging of the initial ANN's predictions over constant minimal sufficient states of  $S_2$ . As the ANN output is discrete, a numerical approximation is necessary for the integral calculations as described in Section 2.2. First, a new data set is created for the discretization, here called the filter data set. It comprises all possible combinations of the three strain components in the range of  $[-0.1, +0.1]$  (same as the test set) with a discretization of  $\Delta\epsilon = 0.0005$  (half of the test set). While finer discretizations provide a better approximation of Eq. (4), the values already reached the Random-Access Memory limit (without consideration of more sophisticated, memory saving programming) causing significantly higher computational cost compared to the following two methods. The  $\{\|\epsilon_{\text{dev}}\|, \text{tr}(\epsilon)\}$ -domain of the minimal sufficient set is moreover discretized by a rectilinear grid in ranges of  $\|\epsilon_{\text{dev}}\| \in [0, 0.2]$  and  $\text{tr}(\epsilon) \in [-0.2, 0.2]$  with discretization steps of  $\Delta\|\epsilon_{\text{dev}}\| = 0.001$  and  $\Delta\text{tr}(\epsilon) = 0.001$ . With that at hand, new improved predictions were calculated according to Eq. (4) for the discretized domain and evaluated for the test set. Interpolation for empty areas was not required in this specific case.

As a second option, Rao-Blackwellization dictates the possibility of data augmentation as introduced in Section 2.4, to employ the knowledge that the minimal sufficient set suffices. It allows to chose one strain component independently while the other two are fixed for given values of  $\|\epsilon_{\text{dev}}\|$  and  $\text{tr}(\epsilon)$ . Theoretically, infinitely many data points could be augmented in the initial strain space with this method. In practice, the amount should be limited in order to keep the training time reasonable. Therefore we apply the following procedure for each initial training data set:

1. Loop over each data point  $i$  in the set and compute  $\|\epsilon_{\text{dev}}\|$  and  $\text{tr}(\epsilon)$
2. Loop over all strain components  $j$  (from  $\epsilon_{11}$ ,  $\epsilon_{22}$  and  $\epsilon_{12}$ )
3. Loop over all elements  $k$  in list  $[-0.1, -0.075, -0.05, -0.025, 0, 0.025, 0.05, 0.075, 0.1]$
4. Set strain component  $j$  to list element  $k$  and compute both other strain components such that the strain state takes the values  $\|\epsilon_{\text{dev}}\|$  and  $\text{tr}(\epsilon)$  for data point  $i$
5. Assign the energy values of point  $i$  to the augmented data point
6. End all loops
7. Add all augmented data points to the initial data set

The list of strain values in step 4 was chosen in accordance with the smallest initial data set. The new augmented training data sets are illustrated in Fig. 5 (bottom). Based on those training sets, again, ANNs (with the same settings as the initial ANNs) were trained and evaluated on the test set.

The third tested possibility to employ Rao-Blackwellization is to optimize the ANN structure. More specifically, we adapt the input layer and train ANNs which receive the minimal sufficient set  $S_2$  instead of the three strain components. As part of a second variation, we do the same with the non-sufficient set  $S_1$  and the non-minimal sufficient set  $S_3$ . All three variants of ANNs are trained on the same initial training data points as the initial ANNs, only with a different parameterization. The training data points parameterized in  $S_2$  are illustrated in Fig. 5 (bottom).

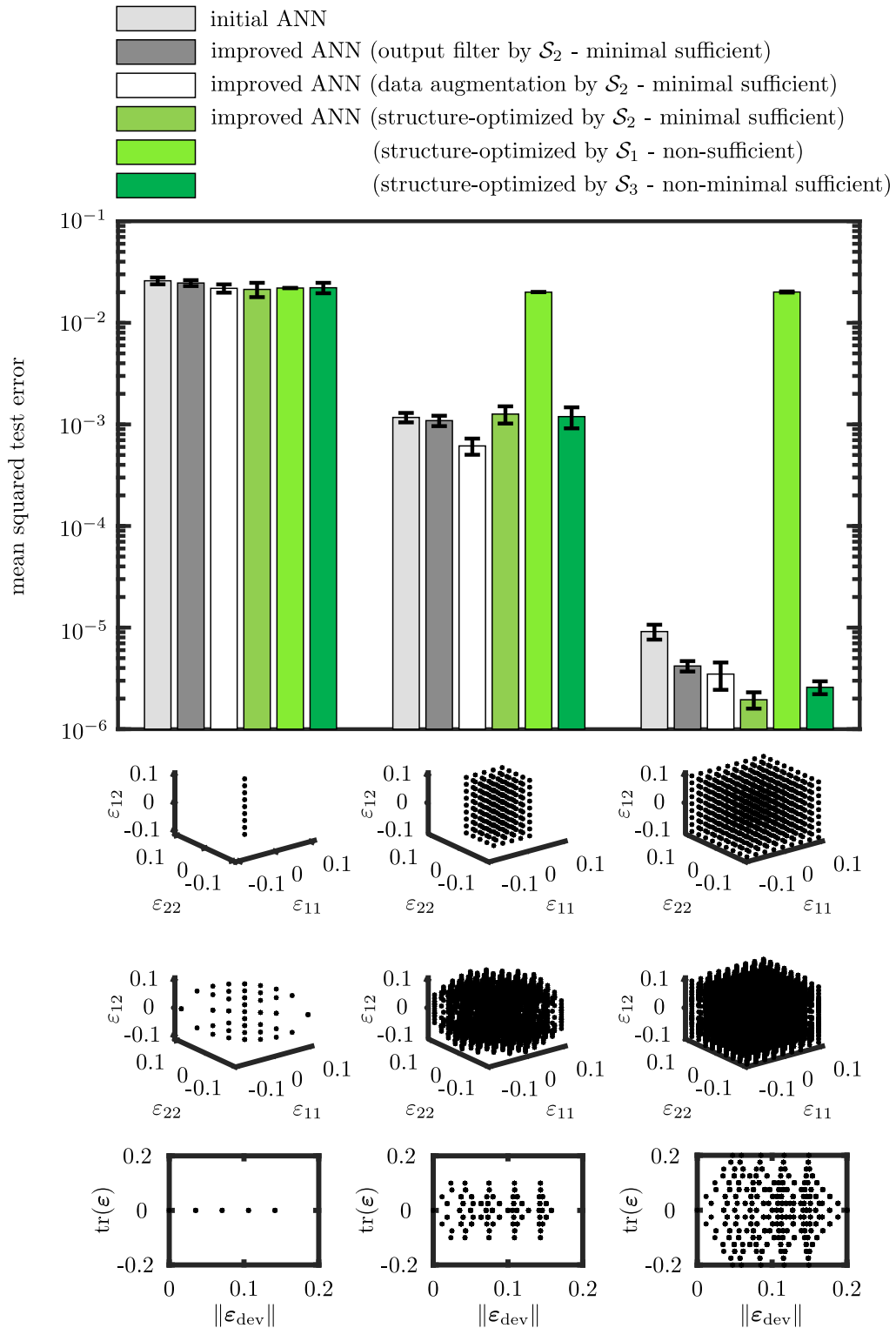


Fig. 5. Comparison of mean squared test errors of the initial, output filtered, data-augmented as well as structure-optimized ANNs. Standard deviations are provided based on five, randomly assigned initial weight distributions. Corresponding training data sets for the initial ANNs are depicted in the upper row, the ones of the data augmentation in the middle and the ones for the ANNs structure-optimized by  $\mathcal{S}_2$  in the lower row.

#### 4.3.2. Results and discussion

Filtering the output and training data augmentation always yields to improvements in terms of the mean squared error, with maximum factors of 2.2 (output filtering) and 2.6 (data augmentation) for the largest training data set, see Fig. 5 (dark gray and white bars compared to light gray bars). Throughout all training data sets, training-data augmentation leads to better results compared to the output filter method. The observed behavior can be explained by the fact that the data augmentation method extrapolates the training data values in the strain component space, thereby covering more of the test data space. Data augmentation is only outperformed by structure optimization of the first layer for the smallest and largest data. For the largest training data set, which has the same parameter range as the test data, the structure-optimized ANN shows a 4.7 times lower error compared to the initial and a 1.8 times lower error compared to the data augmentation ANN.

Two practical limitations emerge along with the generally favorable results of the physics-based Rao–Blackwell scheme. Firstly, the smallest data set is simply too small in order to represent the test data properly. It leads to unsatisfactory predictions for all strategies, despite the observed improvement. Secondly, the structure-optimized ANN trained on the medium-sized data set performs even worse than the initial ANN. This usually contradicts the error inequality of the analytical definition, but the slight deviation can be explained by a poorer numerical extrapolation and possible local minima. Another reason worth mentioning is that the initial input is already a non-minimal sufficient set such that an improvement cannot be guaranteed here.

The comparison of different kinds of sufficiency shows a poor performance for the non-sufficient set  $S_1$ . This result can be expected, but it also underlines quantitatively how well the sufficient set performs compared to a non-sufficient one. For the non-minimal sufficient set  $S_3$ , the error is similar to that of the minimal sufficient set, except for the largest training data set, where the error is slightly higher. One possible explanation for this increased error is that the ANN cannot fully eliminate the influence of the dispensable parameter  $\varepsilon_{11}$ . The presented employment of physics-based Rao–Blackwellization thus clearly helps to efficiently use knowledge about sufficient, but not superfluous, physical information.

In conclusion for the present example of damage energy prediction, physics-based Rao–Blackwellization should be applied via structure optimization if the training data set covers the entire parameter range. It should be used for data augmentation instead, if the training data set covers a smaller parameter range than required by the intended application. Without prior indication, the changes of the proposed scheme range between numerically negligible deviations and clear improvements.

### 5. When physical information can improve ANNs and when it cannot: Experimental example of isotropic elasticity

This example resumes the introductory motivation in Fig. 1, where a rash implementation of physical constraints was shown to worsen a prediction. We now expand the introductory case to a complete ANN prediction of isotropic rubber elasticity trained on DIC data of a highly-elastic 3D-printed TPU (Thermoplastic Polyurethane, Ultimaker 95A) specimen under plane-stress conditions. The principle of utilizing isotropy or homogeneity is already well-known from analytical invariant formulations as well as data-driven predictions, e.g., for PINNs and CANNs [8,13,14]. Here, we will analyze how the employment of isotropy, homogeneity and differentiability follows naturally from physics-based Rao–Blackwellization. At the same time, we will show that the improper employment of auxetic and incompressible limits is not advocated by this strategy and how the modifications perform when comparing numerical predictions with original experimental data.

#### 5.1. Initial data set

The initial state is determined by the raw strain and stress data obtained from tensile stretch experiments (in a Kamrath & Weiss tensile module) with simultaneous picture recording for DIC (via Veddac Strain software), see Fig. 6. Data is recorded for increasing steps of clamp displacement. Piola–Kirchhoff stress values are derived from corresponding force measurements and initial cross-sectional area. Strain data is determined from DIC for 275 subregions evenly distributed in the central region of the specimen. The initial state of the slender and soft specimen showed buckling after mounting via clamps, which was also the initial state for the DIC evaluation. The reference state for the problem at hand was however later set to the zero force level after some stretching. This caused a zero strain level in average but also fluctuations of the DIC strain field, as it refers to a previous reference figure.

Considering analytical model descriptions, this data could be used to fit a hyperelastic isotropic model. Application to an ANN instead requires further adaption. The stress and strain data does not cover the entire domains, e.g., in terms of other load directions. An ANN trained only on the raw data will thus not only perform poorly. It will simply fail, e.g., already prohibiting convergence of mechanical equilibrium when used as a material model in a finite element framework. Among multiple options to adapt the ANN (e.g., see the previous examples improving ANNs in terms of training data, structure or output filters), this example will apply physics-based Rao–Blackwellization directly to the experimental data to better discuss the assumptions employed to the model.

#### 5.2. Data augmentation by sufficient and other physical knowledge

A sufficient information set depends on the respective model (both physical or statistical) and can only be obtained by knowledge about the underlying assumptions. Each assumption can then yield constraints of the ANN. As shown above, the constraints can be realized via Rao–Blackwellization in different ways such as constraining the output with a filter or constraining the input layer. This is in accordance with other technical interpretations of ANN constraints, e.g., in the form of penalty terms [8,12]. The focus of the following investigation will be put more on the physics and thus on the model assumptions and less on the technical constraints as analyzed in the previous example.

To prepare the improvement step of physics-based Rao–Blackwellization, we want to examine six assumptions for the further discussion:

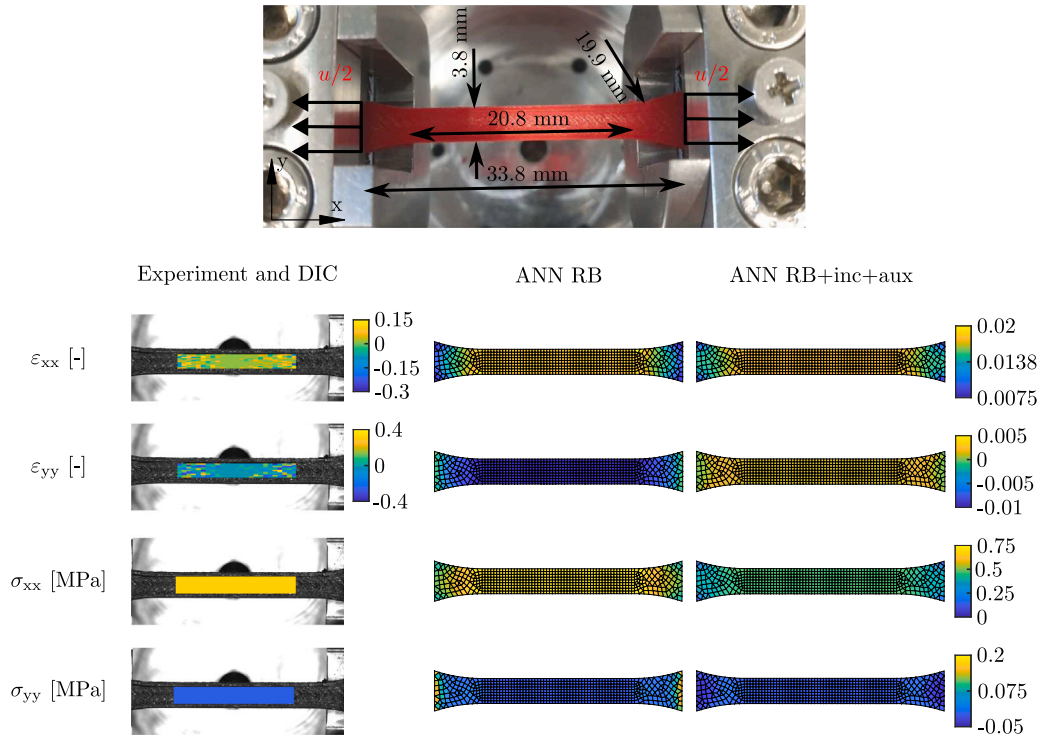


Fig. 6. Experimental set-up (top) and comparison of stress and strain fields determined in the experiment as well as in the finite element simulations for  $u = 0.577$  mm clamp displacement (bottom). Experimental strain fields are obtained from DIC and stress values are calculated from force and area measurements with free lateral boundary conditions, respectively.

1. A homogeneous stress state sufficiently determines a homogeneous strain state and vice versa for each deformation state,
2. the isotropic material dictates material invariance under rotations,
3. the stress–strain relations are continuously differentiable,
4. the loading state remains plane-stress,
5. the material is not incompressible,
6. the material is not auxetic.

It should be emphasized that modeling assumptions naturally show deviations from the real system, some of which can be obvious and some less obvious. In the present example, the force measurements are almost tacitly assigned to stress values inside the central region of the specimen. Strains are however recorded via DIC and show heterogeneous fluctuations that can arise from material imperfections but also from the evaluation algorithm and limited pixel resolution, for instance. Such incoherent data structure can pose a practical challenge for the questions raised in the motivation: what information to incorporate, when to do it and how to do it correctly? Physics-based Rao-Blackwellization provides a reliable answer for those assumptions that yield sufficient information.

### 5.3. Improved ANN material models and incompatible model assumptions

#### 5.3.1. Employing model assumptions

The first model assumption of the above list implies that the (assumed) homogeneous stress state must fully determine the (equally assumed homogeneous) strain state in the central region of the specimen. The specimen shape was accordingly chosen to provide a homogeneous field until the onset of necking, which is not assumed nor observed for the present load regimes. Fluctuations of the DIC-based strain data nevertheless occur and are treated by physics-based Rao-Blackwellization. It dictates for every deformation step, that the sufficient stress value is assigned to a strain value determined from the conditionally averaged-based estimator as defined in (1). Here, the conditional average relates to the condition of a stress level and thus becomes equal to a simple strain average for each load state. This yields to a unique stress–strain path, for example, cf. Fig. 7.

The second model assumption ensures isotropic material behavior and can be used for data augmentation, because stresses and strains sufficiently determine each other, modulo rotations. In the sense of physics-based Rao-Blackwellization, new stress–strain data pairs  $\{\sigma^*, \epsilon^*\}$  are thus generated by rotating a measured pair  $\{\sigma, \epsilon\}$  by rotation matrix  $\mathbf{R}$  as

$$\epsilon^* = \mathbf{R}^T(\vartheta) \cdot \epsilon \cdot \mathbf{R}(\vartheta), \tag{23}$$

$$\sigma^* = \mathbf{R}^T(\vartheta) \cdot \sigma \cdot \mathbf{R}(\vartheta). \quad (24)$$

The rotation angle  $\vartheta$  ranges from 0 to  $2\pi - \Delta\vartheta$  in discrete steps of  $\Delta\vartheta = \pi/18$ . Realizing the Rao–Blackwell scheme via data augmentation is preferred here and in the next step, because it was superior in the aforementioned example for extrapolation tasks of a limited starting data set. Other alternatives such as invariant-based formulations are possible for elastic [45] and inelastic materials [46] as also used in the previous examples. The data augmentation of strain values is nevertheless chosen here to better highlight the effect of data processing in detail. This can be particularly relevant for ANNs that are closer to direct experimental observations [14] and which may not allow a retrospective filter of input or output for their improvement.

The third model assumption does not only guarantee a continuous tangent stiffness. Physics-based Rao-Blackwellization implies that the elastic tangent in the tension regime sufficiently also determines the behavior under compression around the origin. Compression states are not substantially triggered, neither for training data nor for its evaluation. They are thus obtained, up to 1.016% compression strain, by flipping tension states via linear regression. This expansion remains in the linear regime and fully covers the compressive states that appear during the simulation tests.

The fourth model assumption allows to neglect out-of-plane contributions of the stress tensor as information. Accordingly, the transverse contraction in the third direction is also implicitly ignored by only resolving and predicting the two-dimensional state by design of the ANN structure. A three-dimensional extension could be obtained by the same procedure and rotations in the full three-dimensional space, but the comparison to experimental DIC data is facilitated for the plane stress case.

The fifth and sixth model assumptions of the above list cannot be applied to the original data in the present context of physics-based Rao-Blackwellization, because they cannot be utilized for sufficient information transfer. This should of course not be understood as an evaluation of their importance or relevance. These two assumptions are physically sound and should be considered for the solution of mechanical problems with standard materials. Their incompatibility yet indicates that their incorporation into the Rao–Blackwell scheme cannot guarantee an improvement, which was already indicated by the introductory example and will be examined quantitatively now.

A combination of all the above assumptions is hence not suggested by physics-based Rao-Blackwellization. We will thus compare two improvements of the data set for ANN training. The first data set is improved by assumptions 1–4 and shall simply be labeled as “RB”. The second data set additionally considers assumptions 5 and 6 by removing unphysical data and shall be labeled as “RB+inc+aux”.

### 5.3.2. ANN-based finite element simulation compared to experimental reference

The improved data sets are used to train an ANN, “RB” and “RB+inc+aux” respectively, which serves the role of a material model routine in finite element simulations. Note that the raw data set would not yield any feasible result due to the severe lack of training data in the stress and strain domains. The crude initial estimator, in the sense of physics-based Rao-Blackwellization, thus simply fails already by not being able to perform a practicable finite element simulation. The ANNs have a network size of 3-10-10-10-3 with a ReLu activation function and are trained via the SGD method for various epochs with adaptive learning rates and batch sizes. Input of the ANNs are the strain values  $\epsilon_{xx}$ ,  $\epsilon_{yy}$  and  $\epsilon_{xy}$  (due to the symmetry of the strain and stress tensor both respective, off-diagonal components are equal). The training process yields functional ANNs on the material point level for the finite element simulations. The original experimental setting is finally simulated for evaluation of the ANN performance. Therefore, the specimen is meshed via triangular and quadrilateral elements with linear ansatz functions. Boundary conditions are applied according to the depicted boundary value problem, see Fig. 6.

Interestingly, both ANNs yield very similar longitudinal strain ( $\epsilon_{xx}$ ) predictions. These also match the experimental results well. Also the lateral stress in the central specimen region is similar between the ANNs. Both adopted well to the stress-free boundary condition that is also already represented by the original data. A first noticeable difference appears at the clamped ends, though. The “RB+inc+aux” ANN, additionally disregarding incompressible and auxetic limits, predicts neutral to compressive instead of more realistic tensile conditions.

The strongest deviation between both ANNs, “RB” and “RB+inc+aux”, appears for longitudinal stress and lateral strain. The first is already indicated by the stress–strain curve in Fig. 7. Examining the raw data in detail, indicates a systematic error due to limiting DIC resolution. It shifts the strain origin for individual pixels while maintaining the overall course correctly. But again, we want to mimic a scenario that does not allow a retrospective correction of the experimental setup and focus in the present work on improvement with physics-based information instead. While the qualitative evolution is still captured by both ANNs, only the “RB” case based on clean physics-based Rao-Blackwellization allows for a reliable quantitative prediction of the stress–strain relationships. Omitting incompressible and auxetic states almost halves the elastic tangent in some strain states, caused by the strain bias as already discussed in the introduction. The lateral strain shows another severe difference between the ANNs, because insensitive training data becomes more relevant and depends even more on reliable data processing, cf. Fig. 6. To be more specific in terms of the  $-\epsilon_{yy}/\epsilon_{xx}$  ratio, the “RB+inc+aux” case shows a deviation of 119% from the experimental data for 0.577 mm clamp displacement. The deviation of the pure improvement case “RB” is only 7.25% and hence 16 times smaller.

Eventually, these results show that the physics-based Rao-Blackwellization can be successfully used with model assumptions that allow to transfer sufficient information into a finite-element framework. The original experiment is numerically well captured when relying on the pure improvement algorithm. It cannot incorporate all assumptions such as incompressible and auxetic limits, yet it clearly indicates whether such information is not suitable and should be left out — guaranteeing improvement even for crude initial estimators or data availability.

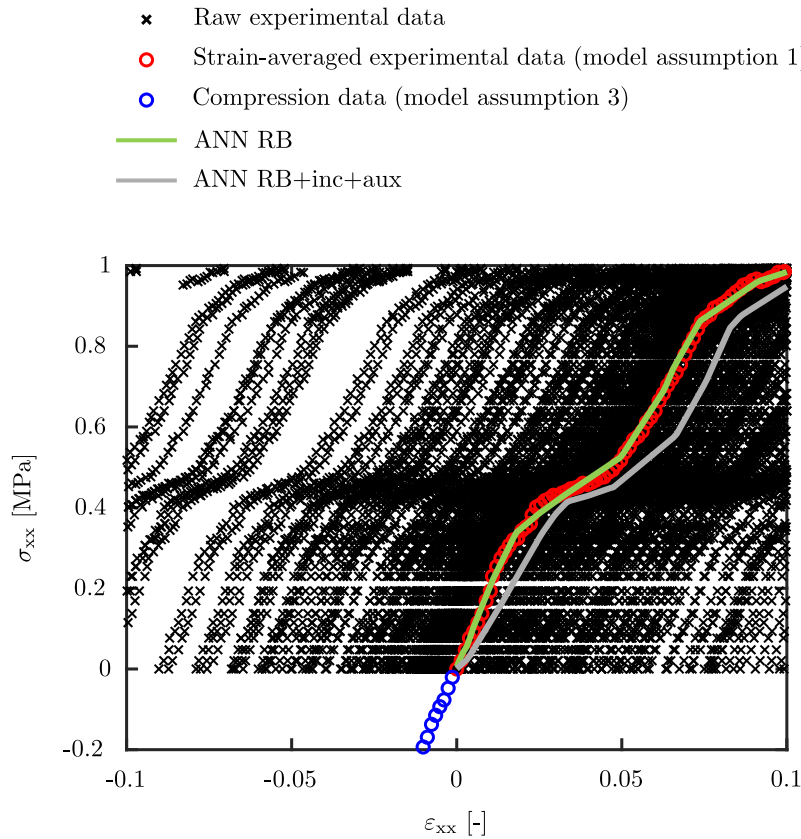


Fig. 7. Raw stress–strain data under uniaxial tension from experimental data and application of assumptions 1 (homogeneity) and 3 (differentiability) and curves predicted by the ANNs trained with different model assumptions, RB vs. RB+inc+aux.

## 6. Conclusion

The concept of Rao-Blackwellization was adapted from its statistical origin to deterministic physical problems. It has been successfully used to improve ANN predictions in the context of material modeling. The basic idea is to unify an ANN’s prediction over sufficient information sets. Identifying sufficient information is the key input for this strategy. It can be obtained from various sources for the problem at hand, for instance: parameters of governing equations fixing the solution, fundamental laws, assumptions like isotropic or quasi-static problems and dimensional analysis. The improvement scheme can then be employed into the ANN by output filtering, training data augmentation or structure adaptation.

Complementing existing approaches to incorporate physical knowledge, the proposed strategy provides an error norm and a guarantee in the form of an error inequality. The presented examples showed various opportunities for the improvement of ANNs. The classification example of an elasto-plastic predictor demonstrated how the scheme suppresses noise and overfitting by post-processing of the output. The example of quasi-brittle damage underlined why sufficient information is required for guaranteed improvement and compared output filtering to data augmentation and structure optimization. Finally, the DIC measurements of flexible, 3D-printed TPU demonstrated that isotropy, homogeneity and differentiability can support the ANN training. The proposed strategy can thus support intuitive as well as less intuitive modeling assumptions. It was shown at the same time, though, that even sound physical assumptions can lead to poorer predictions if they are not supported by a clear improvement scheme such as Rao-Blackwellization.

The present strategy thus complements the scope of existing approaches such as PINNs or CANNs by focusing on the information used, not on the form of the solution. As a first relevant feature, it is guaranteed not to worsen the initial prediction. Secondly, it can improve even crude initial estimators and is thus suitable for situations with an incomplete data basis. Thirdly and more practically, it can be applied flexibly to data, structure or integration of the ANN. These results shall also promote further exchange between statistical modeling and material modeling. Follow-up studies will aim at mathematical aspects such as the impact of parameterization on individual error weighting and numerical inaccuracies. Future work will moreover apply the scheme to information transfer in multi-scale problems, for example, obtaining sufficient information from homogenization of mechanical systems and scale separation [3,47–50]. From a statistical perspective, a more comprehensive transfer in the sense of the Lehmann–Scheffé theorem [22] and the consideration of statistical distributions instead of fixed deterministic model parameters can constitute promising future extensions in the context of model discovery and uncertainty analysis.

## CRediT authorship contribution statement

**Gian-Luca Geuken:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jörn Mosler:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Funding acquisition, Formal analysis, Conceptualization. **Patrick Kurzeja:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgments

Financial support from the German Research Foundation (DFG) via SFB/TRR 188 (278868966), project C01, is gratefully acknowledged. Furthermore, the authors gratefully acknowledge the computing time provided on the Linux HPC cluster at Technical University Dortmund (LiDO3), partially funded in the course of the Large-Scale Equipment Initiative by the German Research Foundation (DFG) as project 271512359.

## Appendix A. Proving how the new prediction never gets worse

The following proof is not fully rigorous in a mathematical sense as some properties are tacitly assumed. Most importantly, we assume that there is a relationship (e.g., between stress and yield behavior) and that integrals are well defined. More information about such peculiarities can be found in [21–24]. These assumptions are however weak when considering well-defined material models with finite deformation energies, limited deformation ranges and rates. The key idea of the following proof closely follows the original Rao–Blackwell theorem and is based on a split of the parameterization into sufficient parameters and remaining parameters not affecting the unknown of interest.

Suppose the correct model prediction is given by  $\theta_*$ , then the error of the initial prediction is defined as

$$\delta_0 = \frac{1}{|\Omega|} \int_{\Omega} (\theta_0 - \theta_*)^2 d\omega. \quad (\text{A.1})$$

This can be rewritten as

$$\frac{1}{|\Omega|} \int_{\Omega} (\theta_0 - \theta_*)^2 d\omega = \frac{1}{|\Omega|} \int_{\Omega} (\theta_0 - \theta_1 + \theta_1 - \theta_*)^2 d\omega \quad (\text{A.2})$$

using the improved model  $\theta_1$  and with the binomial theorem

$$\frac{1}{|\Omega|} \int_{\Omega} (\theta_0 - \theta_*)^2 d\omega = \frac{1}{|\Omega|} \int_{\Omega} (\theta_0 - \theta_1)^2 d\omega + \frac{1}{|\Omega|} \int_{\Omega} (\theta_1 - \theta_*)^2 d\omega + \frac{2}{|\Omega|} \int_{\Omega} (\theta_0 - \theta_1)(\theta_1 - \theta_*) d\omega \quad (\text{A.3})$$

follows. To show that the last term of the right-hand side of (A.3) vanishes, the integral over the domain  $\Omega$  is split into two parts. One integrating over the domain  $\Omega_s$ , where  $S(\omega) = s$  is constant, and another integrating over the remaining parameterization of the physical state  $\Omega_r$ . This leads to

$$\frac{1}{|\Omega|} \int_{\Omega} (\theta_0 - \theta_1)(\theta_1 - \theta_*) d\omega = \frac{1}{|\Omega_r|} \int_{\Omega_r} \frac{1}{|\Omega_s|} \int_{\Omega_s} (\theta_0 - \theta_1)(\theta_1 - \theta_*) d\omega ds \quad (\text{A.4})$$

and, because  $\theta_1$  is constant in  $\Omega_s$ ,

$$\frac{1}{|\Omega|} \int_{\Omega} (\theta_0 - \theta_1)(\theta_1 - \theta_*) d\omega = \frac{1}{|\Omega_r|} \int_{\Omega_r} (\theta_1 - \theta_*) \frac{1}{|\Omega_s|} \int_{\Omega_s} (\theta_0 - \theta_1) d\omega ds. \quad (\text{A.5})$$

Considering only the inner integral and splitting it up, one obtains

$$\frac{1}{|\Omega_s|} \int_{\Omega_s} (\theta_0 - \theta_1) d\omega = \frac{1}{|\Omega_s|} \int_{\Omega_s} \theta_0 d\omega - \frac{1}{|\Omega_s|} \int_{\Omega_s} \theta_1 d\omega. \quad (\text{A.6})$$

This integral vanishes, since

$$\frac{1}{|\Omega_s|} \int_{\Omega_s} \theta_0 d\omega - \frac{\theta_1}{|\Omega_s|} \int_{\Omega_s} 1 d\omega = \theta_0 - \theta_1 = 0 \quad (\text{A.7})$$

**Table B.1**

ANN hyperparameters for the different examples.

Example	Network size	Activation function	Training algorithm	Learning rate	Batch size
Elasto-plastic predictor	2-5-1 2-10-5-1 2-200-200-60-1	Tanh	SGD	Adaptive	Adaptive
Quasi-brittle damage	1/2/3-50-50-50-20-2	ReLU <sup>a</sup>	SGD	0.01	1
Isotropic elasticity	3-10-10-10-3	ReLU <sup>a</sup>	SGD	Adaptive	Adaptive

<sup>a</sup> Last layer has a linear activation function.

by definition of  $\theta_1$ . Now, resubstitution of (A.7) in Eqs. (A.5) and (A.4) and solving (A.3) for the error of the improved model  $\theta_1$  finally leads to

$$\frac{1}{|\Omega|} \int_{\Omega} (\theta_1 - \theta_{\star})^2 d\omega = \frac{1}{|\Omega|} \int_{\Omega} (\theta_0 - \theta_{\star})^2 d\omega - \underbrace{\frac{1}{|\Omega|} \int_{\Omega} (\theta_0 - \theta_1)^2 d\omega}_{\geq 0}. \quad (\text{A.8})$$

Since the second term of the right-hand side is always greater or equal to zero, it is confirmed that  $\theta_1$  results in a lower or equal error compared to the one of the initial model  $\theta_0$

$$\frac{1}{|\Omega|} \int_{\Omega} (\theta_1 - \theta_{\star})^2 d\omega \leq \frac{1}{|\Omega|} \int_{\Omega} (\theta_0 - \theta_{\star})^2 d\omega \quad (\text{A.9})$$

and the proof is complete. Also note that this scheme is idempotent and another Rao-Blackwellization of  $\theta_1$  does not lead to further improvement  $\theta_2$ , since

$$\theta_2 = \mathbb{A}(\theta_1 | S = s) = \frac{1}{|\Omega_s|} \int_{\Omega_s} \theta_1 d\omega = \theta_1. \quad (\text{A.10})$$

## Appendix B. Overview of ANN hyperparameters

Table B.1 summarizes the ANN hyperparameters for every example. They were chosen according to convergence studies and by taking into account problem specific requirements. For the elasto-plastic predictor example the Tanh delivers a suitable activation function due to its bounded output. For the two other examples the common ReLU function provided the best result when combined with a linear activation of the last layer.

## References

- [1] F. Brezzi, M. Fortin (Eds.), Mixed and Hybrid Finite Element Methods, in: Springer Series in Computational Mathematics, vol. 15, Springer, New York, 1991, <http://dx.doi.org/10.1007/978-1-4612-3172-1>.
- [2] T. Belytschko, W.K. Liu, B. Moran, K.I. Elkhodary, *Nonlinear Finite Elements for Continua and Structures*, second ed., Wiley, 2014.
- [3] P. Kurzeja, The criterion of subscale sufficiency and its application to the relationship between static capillary pressure, saturation and interfacial areas, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 472 (2016) 20150869, <http://dx.doi.org/10.1098/rspa.2015.0869>.
- [4] U. Forssell, P. Lindskog, Combining semi-physical and neural network modeling: An example of usefulness, IFAC Proc. Vol. 30 (1997) 767–770, [http://dx.doi.org/10.1016/S1474-6670\(17\)42938-7](http://dx.doi.org/10.1016/S1474-6670(17)42938-7).
- [5] G. Liang, K. Chandrashekhara, Neural network based constitutive model for elastomeric foams, Eng. Struct. 30 (2008) 2002–2011, <http://dx.doi.org/10.1016/j.engstruct.2007.12.021>.
- [6] H.K.D.H. Bhadeshia, Neural networks and information in materials science, Statist. Anal. Data Min.: ASA Data Sci. J. 1 (2009) 296–305, <http://dx.doi.org/10.1002/sam.10018>.
- [7] I. Lagaris, A. Likas, D. Fotiadis, Artificial neural networks for solving ordinary and partial differential equations, IEEE Trans. Neural Netw. 9 (5) (1998) 987–1000, <http://dx.doi.org/10.1109/72.712178>.
- [8] M. Raissi, P. Perdikaris, G. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. 378 (2019) 686–707, <http://dx.doi.org/10.1016/j.jcp.2018.10.045>.
- [9] A. Kashеfi, T. Mukerji, Physics-informed pointnet: A deep learning solver for steady-state incompressible flows and thermal fields on multiple sets of irregular geometries, J. Comput. Phys. 468 (2022) 111510, <http://dx.doi.org/10.1016/j.jcp.2022.111510>.
- [10] A. Jagtap, G. Karniadakis, Extended physics-informed neural networks (xpinns): A generalized space–time domain decomposition based deep learning framework for nonlinear partial differential equations, Commun. Comput. Phys. 28 (2020) 2002–2041, <http://dx.doi.org/10.4208/cicp.OA-2020-0164>.
- [11] A. Henkes, H. Wessels, R. Mahnken, Physics informed neural networks for continuum micromechanics, Comput. Methods Appl. Mech. Engrg. 393 (2022) 114790, <http://dx.doi.org/10.1016/j.cma.2022.114790>.
- [12] S. Cuomo, V.S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, F. Piccialli, Scientific machine learning through physics-informed neural networks: Where we are and what's next, J. Sci. Comput. 92 (2022) 88, <http://dx.doi.org/10.1007/s10915-022-01939-z>.
- [13] K. Linka, M. Hillgärtner, K.P. Abdolazizi, R.C. Aydin, M. Itskov, C.J. Cyron, Constitutive artificial neural networks: A fast and general approach to predictive data-driven constitutive modeling by deep learning, J. Comput. Phys. 429 (2021) 110010, <http://dx.doi.org/10.1016/j.jcp.2020.110010>.
- [14] S. Tang, G. Zhang, H. Yang, Y. Li, W.K. Liu, X. Guo, Map123: A data-driven approach to use 1d data for 3d nonlinear elastic materials modeling, Comput. Methods Appl. Mech. Engrg. 357 (2019) 112587, <http://dx.doi.org/10.1016/j.cma.2019.112587>.
- [15] T. Kirchdoerfer, M. Ortiz, Data-driven computational mechanics, Comput. Methods Appl. Mech. Engrg. 304 (2016) 81–101, <http://dx.doi.org/10.1016/j.cma.2016.02.001>.

- [16] A. Schowtjak, R. Schulte, T. Clausmeyer, R. Ostwald, A.E. Tekkaya, A. Menzel, Adapt — a diversely applicable parameter identification tool: Overview and full-field application examples, *Int. J. Mech. Sci.* 213 (2022) 106840, <http://dx.doi.org/10.1016/j.ijmecsci.2021.106840>.
- [17] J.F. Unger, C. Könke, Neural networks as material models within a multiscale approach, *Comput. Struct.* 87 (2009) 1177–1186, <http://dx.doi.org/10.1016/j.compstruc.2008.12.003>.
- [18] A. Waseem, T. Heuzé, M.G. Geers, V.G. Kouznetsova, L. Stainier, Data-driven reduced homogenization for transient diffusion problems with emergent history effects, *Comput. Methods Appl. Mech. Engrg.* 380 (2021) 113773, <http://dx.doi.org/10.1016/j.cma.2021.113773>.
- [19] C. Settigast, G. Hütter, M. Kuna, M. Abendroth, A hybrid approach to simulate the homogenized irreversible elastic–plastic deformations and damage of foams by neural networks, *Int. J. Plast.* 126 (2020) 102624, <http://dx.doi.org/10.1016/j.ijplas.2019.11.003>.
- [20] S. Freitag, B. Cao, J. Ninić, G. Meschke, Recurrent neural networks and proper orthogonal decomposition with interval data for real-time predictions of mechanised tunnelling processes, *Comput. Struct.* 207 (2018) 258–273, <http://dx.doi.org/10.1016/j.compstruc.2017.03.020>.
- [21] D. Blackwell, Conditional expectation and unbiased sequential estimation, *Ann. Math. Stat.* 18 (1947) 105–110, <http://dx.doi.org/10.1214/aoms/1177730497>.
- [22] E.L. Lehmann, G. Casella, *Theory of Point Estimation*, second ed., Springer-Verlag, 1998.
- [23] C. Rao, *Linear Statistical Inference and its Applications*, in: *Wiley Series in Probability and Statistics - Applied Probability and Statistics Section Series*, Wiley-Interscience, 2001.
- [24] P. Bickel, K. Doksum, *Mathematical Statistics, Basic Ideas and Selected Topics*, second ed., vol. 1, Prentice Hall, 2006.
- [25] A. Doucet, N.d. Freitas, K.P. Murphy, S.J. Russell, Rao-blackwellised particle filtering for dynamic bayesian networks, in: *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence*, UAI'00, 2000, pp. 176–183, <http://dx.doi.org/10.48550/arXiv.1301.3853>.
- [26] D.J.C. MacKay, *Information Theory, Inference and Learning Algorithms*, Cambridge University Press, 2003.
- [27] V.L. Deringer, A.P. Bartók, N. Bernstein, D.M. Wilkins, M. Ceriotti, G. Csányi, Gaussian process regression for materials and molecules, *Chem. Rev.* 121 (16) (2021) 10073–10141, <http://dx.doi.org/10.1021/acs.chemrev.1c00022>.
- [28] B. Chen, L. Shen, H. Zhang, Gaussian process regression-based material model for stochastic structural analysis, *ASCE-ASME J. Risk Uncertain. Eng. Syst.* A 7 (3) (2021) 04021025, <http://dx.doi.org/10.1061/AJRUA6.0001138>.
- [29] M. Böttinger, M. Faes, A. Menzel, M. Valdebenito, Effect of uncertainty of material parameters on stress triaxiality and lode angle in finite elasto-plasticity—a variance-based global sensitivity analysis, *Adv. Ind. Manuf. Eng.* 7 (2023) 100128, <http://dx.doi.org/10.1016/j.aime.2023.100128>.
- [30] Y. Wang, N. Wagner, J.M. Rondinelli, Symbolic regression in materials science, *MRS Commun.* 9 (3) (2019) 793–805, <http://dx.doi.org/10.1557/mrc.2019.85>.
- [31] E. Kablman, A.H. Kolody, J. Kronsteiner, M. Kommenda, G. Kronberger, Application of symbolic regression for constitutive modeling of plastic deformation, *Appl. Eng. Sci.* 6 (2021) 100052, <http://dx.doi.org/10.1016/j.applsc.2021.100052>.
- [32] F. Santosa, W.W. Symes, Linear inversion of band-limited reflection seismograms, *SIAM J. Sci. Stat. Comput.* 7 (4) (1986) 1307–1330, <http://dx.doi.org/10.1137/0907087>.
- [33] R. Tibshirani, Regression shrinkage and selection via the lasso, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 58 (1) (1996) 267–288, <http://www.jstor.org/stable/2346178>.
- [34] R.A. Fisher, E.J. Russell, On the mathematical foundations of theoretical statistics, *Philos. Trans. R. Soc. Lond. Ser. A* 222 (1922) 309–368, <http://dx.doi.org/10.1098/rsta.1922.0009>.
- [35] C. Miehe, S. Göktepe, F. Lulei, A micro-macro approach to rubber-like materials—part I: the non-affine micro-sphere model of rubber elasticity, *J. Mech. Phys. Solids* 52 (2004) 2617–2660, <http://dx.doi.org/10.1016/j.jmps.2004.03.011>.
- [36] E.M. Arruda, M.C. Boyce, A three-dimensional constitutive model for the large stretch behavior of rubber elastic materials, *J. Mech. Phys. Solids* 41 (1993) 389–412, [http://dx.doi.org/10.1016/0022-5096\(93\)90013-6](http://dx.doi.org/10.1016/0022-5096(93)90013-6).
- [37] G.-L. Geuken, J. Mosler, P. Kurzeja, Optimizing artificial neural networks for mechanical problems by physics-based rao-blackwellization: Example of a hyperelastic microsphere model, *PAMM* 22 (1) (2023) e202200325, <http://dx.doi.org/10.1002/pamm.202200325>.
- [38] P. Kurzeja, C. Sievers, J. Mosler, Improving constitutive equations in multiscale modelling by means of the sufficiency criterion using the example of nano wire contraction, *Proc. Appl. Math. Mech.* 18 (2018) e201800292, <http://dx.doi.org/10.1002/pamm.201800292>.
- [39] D.K. Klein, M. Fernández, R.J. Martin, P. Neff, O. Weeger, Polyconvex anisotropic hyperelasticity with neural networks, *J. Mech. Phys. Solids* 159 (2022) 104703, <http://dx.doi.org/10.1016/j.jmps.2021.104703>.
- [40] M. Fernández, M. Jamshidian, T. Böhlke, K. Kersting, O. Weeger, Anisotropic hyperelastic constitutive models for finite deformations combining material theory and data-driven approaches with application to cubic lattice metamaterials, *Comput. Mech.* 67 (2021) 653–677, <http://dx.doi.org/10.1007/s00466-020-01954-7>.
- [41] L. De Lorenzis, C. Maurini, Nucleation under multi-axial loading in variational phase-field models of brittle fracture, *Int. J. Fract.* 237 (2021) 61–81, <http://dx.doi.org/10.1007/s10704-021-00555-6>.
- [42] K. Langenfeld, P. Kurzeja, J. Mosler, How regularization concepts interfere with (quasi-)brittle damage: a comparison based on a unified variational framework, *Contin. Mech. Thermodyn.* 34 (2022) 1517–1544, <http://dx.doi.org/10.1007/s00161-022-01143-2>.
- [43] G. Del Piero, D.R. Owen, Structured deformations of continua, *Arch. Ration. Mech. Anal.* 124 (1993) 99–155, <http://dx.doi.org/10.1007/BF00375133>.
- [44] B. Schmidt, F. Fraternali, M. Ortiz, Eigenfracture: An eigendeformation approach to variational fracture, *Multiscale Model. Simul.* 7 (2009) 1237–1266, <http://dx.doi.org/10.1137/080712568>.
- [45] L. Linden, D.K. Klein, K.A. Kalina, J. Brummund, O. Weeger, M. Kästner, Neural networks meet hyperelasticity: A guide to enforcing physics, *J. Mech. Phys. Solids* 179 (2023) 105363, <http://dx.doi.org/10.1016/j.jmps.2023.105363>.
- [46] K.A. Meyer, F. Ekke, Thermodynamically consistent neural network plasticity modeling and discovery of evolution laws, *J. Mech. Phys. Solids* 180 (2023) 105416, <http://dx.doi.org/10.1016/j.jmps.2023.105416>.
- [47] M. te Vrugt, S. Hossenfelder, R. Wittkowski, Mori-zwanzig formalism for general relativity: A new approach to the averaging problem, *Phys. Rev. Lett.* 127 (2021) 231101, <http://dx.doi.org/10.1103/PhysRevLett.127.231101>.
- [48] A. Javili, P. Steinmann, J. Mosler, Micro-to-macro transition accounting for general imperfect interfaces, *Comput. Methods Appl. Mech. Engrg.* 317 (2017) 274–317, <http://dx.doi.org/10.1016/j.cma.2016.12.025>.
- [49] T. Heitbreder, N.S. Ottosen, M. Ristinmaa, J. Mosler, On damage modeling of material interfaces: Numerical implementation and computational homogenization, *Comput. Methods Appl. Mech. Engrg.* 337 (2018) 1–27, <http://dx.doi.org/10.1016/j.cma.2018.03.023>.
- [50] J.H. Cushman, L.S. Bennethum, B.X. Hu, A primer on upscaling tools for porous media, *Adv. Water Resour.* 25 (2002) 1043–1067, [http://dx.doi.org/10.1016/S0309-1708\(02\)00047-7](http://dx.doi.org/10.1016/S0309-1708(02)00047-7).