

PERSPECTIVE

Open Access



From mechanism-based to data-driven approaches in materials science

Stefan Hiemer^{1*}  and Stefano Zapperi^{2,3}

*Correspondence:

stefan.hiemer@fau.de

¹Institute of Materials Simulation,
Department of Materials Science
and Engineering,
Friedrich-Alexander-University
Erlangen-Nuremberg, Fürth,
Germany

Full list of author information is
available at the end of the article

Abstract

A time-honored approach in theoretical materials science revolves around the search for basic mechanisms that should incorporate key feature of the phenomenon under investigation. Recent years have witnessed an explosion across areas of science of a data-driven approach fueled by recent advances in machine learning. Here we provide a brief perspective on the strengths and weaknesses of mechanism based and data-driven approaches in the context of the mechanics of materials. We discuss recent literature on dislocation dynamics, atomistic plasticity in glasses focusing on the empirical discovery of governing equations through artificial intelligence. We conclude highlighting the main open issues and suggesting possible improvements and future trajectories in the fields.

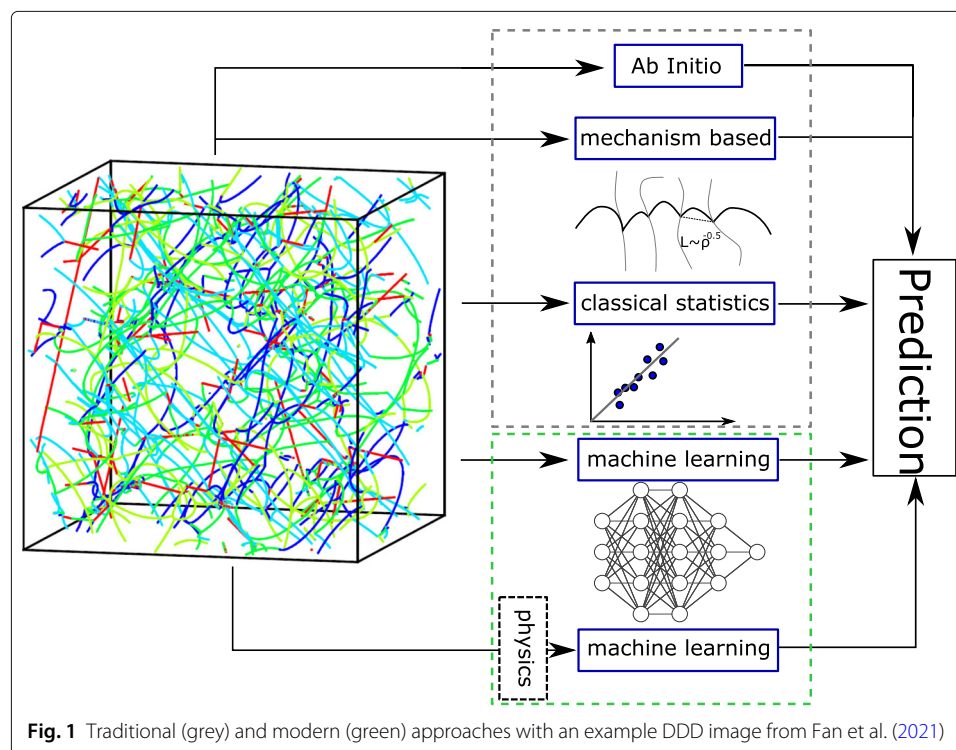
Keywords: Machine learning, Perspective, Dislocation dynamics, Plasticity in glasses, Equation discovery

Introduction

The common goal of continuum mechanics is to establish the macroscopic response to a stimulus in a given geometry for a certain material. This response can be a simple scalar quantity like a force acting on a cantilever or a fully fledged tensorial constitutive model. Ideally, to develop such macroscopic relationships, one would like to derive the governing equations from first principles starting from a microscopic description under a set of approximations. A well established example is the theory of linear elasticity (Landau and Lifshitz 1986), which can be derived based on general symmetry consideration or even as a large-scale limit of an atomic scale description of a crystal. As soon as one considers applications outside the linear regime like in plasticity, a rigorous approach becomes impossible as the approximation from first principles is neither analytically obvious nor computationally feasible. To deal with this issue, traditionally scientists have attempted to derive models from simple underlying material mechanisms, physical analogies or (semi-)empirical considerations which are then corroborated by fitting the usually limited data that were available. Examples of this general approach are provided by dislocation-based theories (Bulatov and Cai 2006; Bertin et al. 2020), spring-damper visco-elastic models (Macosko 1994; Bird et al. 1987) and the different yield hypothesis' in engineering.

The alternative to a mechanism based approach is a data-driven one, which in the past was reduced to either fitting data with trial functions, multivariate or linear combinations of hand-selected functions or linear decomposition techniques like principal component analysis, which in mechanics is known as proper orthogonal decomposition. The first approach suffers inevitably from the fact that humans are not able to visualize data fitting in more than three dimensions, the second is limited by a bias regarding the functional relations and the third performs poorly for non-linear patterns due to its intrinsically linear nature. Modern machine learning (ML) techniques like neural networks bypass these problems as they do not rely on the choice on a proper function thus as a positive side effect taking human bias out of the process. While the success of ML in speech and image processing is common knowledge by now, new data-driven approaches have also outperformed traditional hand-crafted feature methods in computational chemistry as well as linear filtering techniques in computational fluid mechanics. Some notable applications are the structure based prediction of molecular properties (Rupp et al. 2012; Schütt et al. 2017; Gastegger et al. 2020), the development of accurate density functional theory force fields (Behler and Parrinello 2007; Unke et al. 2020; Chmiela et al. 2017) and the discovery of the underlying governing equations of dynamical systems and flow (Champion et al. 2019; Brunton et al. 2016; Rudy et al. 2019).

In this perspective, we share our views on how traditional mechanism based approaches compare to modern methods (Fig. 1) emanating purely from observational data assisted by ML techniques. A crucial aspect that we will underline is the relevance of principles based on physical insight that should be kept in mind when applying ML. We will discuss examples in the fields of dislocations, plastic deformation of glasses and equation discovery, focusing on the limitations of the method and suggesting possible routes for improvement.



Machine learning approaches in physics and mechanics

In this section, we are going to explain the usual workflow of machine learning in general, point to the underlying physical considerations that should be followed and give a short comment on the difference of simulation and experiments. The typical *modus operandi* in ML is to split the available data in training and test sets. One then trains an estimator $f(X, \alpha)$ to predict a target Y with features X extracted from the data by minimizing the cost function C with respect to the trainable parameters α . A model's performance is judged by some summary metric evaluated on the test set which should ideally be dimensionless and avoid error cancellation. The general goal of feature extraction is to obtain a dimensional reduction from the raw data and thus a favourable representation for the prediction, reducing the need of large training sets and producing higher prediction accuracy. Ideally the features X should be chosen following a few guiding principles adapted from approaches in quantum chemistry (Langer et al. 2020): i) the symmetries of the underlying problem must be preserved. In other words, features must be invariant under transformations which leave Y unchanged, ii) The mathematical properties of the problem, like continuity and differentiability with regards to Y , should be preserved by the features. iii) Features should be chosen so that the procedure is computationally efficient. iv) It must be possible to generalize the method to different system sizes so that systems of different sizes could be expressed by the same number of features. v) It should be possible to reconstruct the raw input from X so that the features provide a *lossless representation*.

Criteria i) and iii) should be met at all times while the other conditions could be abandoned in specific cases. Feature construction can be done by hand or by the estimator, as modern convolutional neural networks (CNN) in image recognition do. However, for data not located on a simple grid akin to an image, these well established estimators cannot be employed. Symmetry functions or derived from that continuous filter convolutions however solve that problem for off-grid objects (Schütt et al. 2017; Behler and Parrinello 2007). Regarding the choice of the estimator, researchers should consider their sample size, desired accuracy and interpretability. The sample size narrows the number of possible estimators down due to high sample demand or to poor computation/memory scaling. Accuracy and interpretability can have a trade off relationship as deep learning approaches are very accurate in many applications but remain often quite difficult to interpret. In contrast, decision trees or even linear regression models are much easier to understand than deep learning but are far from optimal with respect to performance (Lapuschkin et al. 2019). For the special case of obtaining interatomic potentials, accuracy is the key, so interpretability is secondary. On the other hand, if we wish to understand a physical process, interpretability is often more important than accuracy.

Machine learning in dislocation dynamics and crystal plasticity

Simulations of dislocations-based plasticity can be performed at different levels in well established spatio-temporal hierarchy, ranging from atomistic molecular dynamics (MD), mesoscale discrete dislocation dynamics (DDD) and continuum dislocation dynamics (CDD) and macroscale crystal plasticity (CP). MD and DDD shed light on the mutual interactions of dislocations and dislocation interactions with defects on a nanometer (MD) and micrometer (DDD) scale while CDD and CP describe dislocation pattern formation, microstructure evolution and mechanical response on dimensions relevant for macroscopic mechanical testing. The hierarchical simulation workflow is to pass

parameters from the lower scale models to higher scale theories, in order to develop models better incorporating the characteristics of a lower scale simulations or to benchmark existing models. Problems arise from the fact that these methods act at different space and time scales and overlapping regions enabling direct benchmarking are either non-existent or small and come at large computational costs. The issue of closure of the hierarchy of equations describing n -point densities in CDD is also not completely resolved. Finally, the analysis of resulting structures and the connection to the mechanical response is also not straightforward.

A series of ML applications to dislocation plasticity has been published by Laurson and coworkers. In their first work, the goal was to predict the stress strain behaviour of individual samples of 2D DDD simulations from internal stresses, encapsulated into statistical moments, magnitudes, Fourier coefficients and some other hand-crafted features. The predictions turned out to be moderately successful ($R^2 \sim 0.5$ for undeformed and $R^2 \sim 0.7$ for prestretched samples) as their estimators failed to predict burst activity which the authors argued to be stochastic by nature, thus limiting the predictability of DDD in general (Salmenjoki et al. 2018). In a more recent work the same authors investigated whether the transition from dislocation pinning to jamming in the presence of spherical precipitates could be detected by the systems dislocation structure. To that end, the system was discretized by voxels of the (spatial) Fourier transform of the time evolution of the geometrically necessary dislocation density, the dislocation spacing correlation and a parameter describing the dislocation junction lengths. To detect the transition, an unsupervised training scheme was employed: Under the assumption that the system described by a parameter p switches phase at a critical value p_c , a binary classifier should maximize its accuracy for phase labels generated with the binarization threshold being the true critical value. By solving this maximization problem, p_c could be found. The training was done independently for each feature. Interestingly the critical precipitate density found for each feature was in reasonable agreement across all three features (Salmenjoki et al. 2020).

A clever way to parametrize the constitutive model for CP from DDD has been shown in Messner et al. (2017). The CP model was developed to describe the slip resistance on a slip system taking into account interaction with all other slip systems by an interaction matrix H . The evolution of the dislocation density was described by another matrix K . Technically H and K have 1521 and 59319 free parameters, but the authors reduce these numbers by taking into account crystal and physical symmetries in the system and applying LASSO regression. In LASSO regression the model chooses a trade-off during training between the number and magnitude of the free parameters and the prediction accuracy on the training set to prevent overfitting (Tibshirani 1996). The number of free parameters for H and K was reduced to 27% and a staggering 0.3% of non-zero components. Another interesting application to extend higher order CP models to larger systems was published recently (Pandey and Pokharel 2020). A number of CP simulations for uniaxial tension with different cell sizes were generated and $3 \times 3 \times 3$ cubes extracted. The goal was to predict the microstructure evolution of the center voxel in the cube over a series of 13 strain steps with a recurrent neural network. This approach has three advantages: i) Each CP simulation yields a high number of training samples thus making dataset generation cheap ii) Each sample is small in terms of memory making the training cheap in

terms of memory consumption iii) The model can be applied to systems of arbitrary size by decomposition and reconstruction of the simulation cell.

The data-driven approaches applied to DDD, apart from a 1D case (Sarvilahti et al. 2020) use hand-crafted features in the spirit of traditional phenomenology, although sophisticated inclusion of crystal symmetries or other physical insights have not yet been used (Salmenjoki et al. 2018; Salmenjoki et al. 2020; Steinberger et al. 2019). While working along known phenomenons is beneficial to support and understand existing models, it is unlikely to reveal a lot of new physics. Symmetry functions from computational chemistry cannot be used as they cannot account for the long range interaction of dislocations, do not include the crystallographic symmetries of the system and the orientation of the dislocation configuration with respect to an external load. The alternative route of mapping the dislocations onto a grid to use convolutional neural networks, as in image recognition, and its generalization to 3D leads to significant information loss due to discretization, exploding memory on reasonably large volumes and neglects crystallographic information. So the search for a way to enable machine crafted features is in our opinion the main issue for ML approaches to DDD. In CP, the route is more straightforward and the two publications discussed above (Messner et al. 2017; Pandey and Pokharel 2020) have already shone light on what can be done. Crystallographic information can be incorporated into the underlying constitutive model via LASSO or other regularization methods. CP simulations can be used as input by off the shelf neural network architectures as they are already located on a grid for most cases. Two interesting questions are i) can nonlocal models be learned without simply increasing the size of the extracted cube as in Pandey and Pokharel (2020)? ii) Can an estimator be trained which predicts global properties like a stress-strain curve of microstructures of different (grid) size?

Data-driven approaches in atomistic glasses

An interesting challenge in amorphous solids is to identify which atoms will plastically deform due to thermally activated processes or under an external load. We focus here on the second task where the loading scenario is always simple shear. Two papers were published by Cubuk, Schoenholz and coworkers (Cubuk et al. 2015; Bapst et al. 2020). The first approach to this problem (Cubuk et al. 2015) was based on the use the symmetry functions to encode the neighbourhood of an atom, as these descriptors are translation and rotation invariant, and to classify whether its non-affine displacement (Falk and Langer 1998) has exceeded a certain threshold in an athermal quasistatic loading scenario. This was done for a range of different thresholds with the best accuracy smaller but close to 75%. The studied glass system was a Kob-Andersen system under shear and the estimators support vector classifiers. Through closer inspection of the symmetry features, deformed and undeformed particles could be distinguished by different distributions in the angular symmetry functions. In their second attempt (Bapst et al. 2020), for the same glass system, the authors tried to predict a simplified non-affine displacement measure via a graph neural network. For that purpose all atoms of one simulation run are embedded as a graph where each node represents an atom and is connected to other atoms within a cutoff radius. After evolving the graph neural network a few times, each node yields the prediction for each individual atom. The success of this model was measured by the Pearson correlation coefficient p

$$p = \frac{\text{Cov}(Y, f(X))}{\sqrt{\text{Var}(Y)\text{Var}(f(X))}}. \quad (1)$$

The numerator is the covariance of the predicted value and the true value. We notice that the covariance allows for errors to cancel. It would therefore be more appropriate to employ more standard metrics and loss functions for regression, such as the coefficient of determination, the mean squared error or the mean absolute error where errors only add up. The use of the Pearson coefficient leads to a too benevolent assessment of a model when compared to the other better designed metrics.

A different approach was taken by Wang et al. (2019) who construct short and medium range features from interstice distributions, Voronoi derived quantities and cluster identifications. Any atoms with a non-affine displacement greater than 5 Å are deemed plastically deformed. The authors investigated a series of metallic glasses at different compositions and quench rates. Accuracy reached in this work goes up to 77% which is a similar performance to Cubuk et al. (2015), but with far fewer features (i.e. 15). Furthermore, the approach seems to generalize reasonably across different compositions and quench rates. In another work by the same group, the authors added additional features to their approach and also investigated the robustness of their descriptors with respect to thermal fluctuations (Wang et al. 2020). In a more recent publication Fan et al. (2021) considered binary metallic glass and binary Lennard Jones systems under different loading protocols. To analyze the systems, the authors created an image for every atom, capturing the neighbourhood of an atom within a cutoff radius. For each pixel, radial symmetry functions for each species are calculated, so that the images have two channels. This procedure obviously sacrifices rotation invariance, which is fixed by rotating each image so that shearing is done along the X and tension/compression along the Y direction. This is a clever thing to do, as the orientation of the local environment with respect to the box strain state is important since local atomic environments are anisotropic. The images created in this way were then fed to a convolutional neural network to classify plastically deformed particles. The results outperformed the graph neural network by Bapst et al. (2020) and support vector machines trained on symmetry functions (Behler and Parrinello 2007).

Apart from Fan's work, no other work has understood that choosing translation and rotation invariant features without consideration of the loading scenario is actually contrary to the symmetries of the underlying atomistic problem. While Fan's workaround is advantageous in terms of simplicity, their method of rotation of the coordinate system works only for simple strain protocols. A generalization to complex loading states should be done especially if one later wants to develop a continuum theory based on the ML findings.

Discovery of functions and partial differential equations from data

Methods for discovering from data the governing equations as a set of algebraic equations is similar to the path for partial differential equations (PDE). These methods were developed in the context of dynamical systems, so the basic mathematical assumption to the model is that we are interested in the time evolution of some quantity $\dot{X} = f(X)$ and look for an approximation of f (Brunton et al. 2016). A collection of possible functions Θ including cross terms is constructed and measurements of X, \dot{X} are performed. By com-

binning the trial functions in a linear combination with the column vector of coefficients Ξ it boils down to a linear problem

$$\dot{X} = \Theta(X)\Xi \quad (2)$$

where $\Theta(X)$ is the matrix of possible functions evaluated with the measured X , where each column represents one trial function. This problem then is solved via a regularization ansatz similar to LASSO enforcing sparsity of the components thus limiting the model to the governing equations (Brunton et al. 2016). The generalization to PDE is straightforward as additional to X, \dot{X} spatial derivatives have to be measured and included in the functional guesses Θ (Rudy et al. 2017; Rudy et al. 2019). Neural network based autoencoder have already been combined with this approach to find the optimal coordinates and equations the solution of problems (Champion et al. 2019). This ansatz however has the downside that the transformation to the discovered coordinate system and back is done via neural networks and the back-transformation loses information.

In our opinion, this branch of ML is particularly interesting for mechanics as a proper inclusion of the symmetries can truly lead to the discovery of new equations. Symmetries can also be discovered from data, by including a functions of different symmetries in the function library Θ and regularizing heavily e. g. via LASSO in order to keep only the most important functions/symmetries. Also contrary to deep learning approaches the discovered equations are interpretable and also it is fairly easy to communicate them to fellow researchers in either publications or personal communication. Extensions to the existing methods is possible, as the developers have made their programs publicly available in a python package (de Silva et al. 2020). The disadvantage of this method is, that it is only as clever as the researcher employing it since this person decides which functions enter Θ . Possible applications can be the discovery of governing equations for CDD and CP or the development of continuum models for the new class of meta-materials.

Dealing with experimental data

As most of this manuscript deals with simulations, we will highlight shortly the major obstacles in applying the same methods to experiments. Training models on extracted scalar quantities like composition-property dependence is completely analogue to simulation and straight forward. Drawing conclusion from 2D/3D data is, however, more difficult. In the case of images one has to deal with a few issues: Parts of the information are limited by the precision of the measurement and noise, while data about the third dimension is completely missing. Apart from ultra-thin cases the missing dimension will affect the prediction and in some cases render it entirely impossible. In the case of the model being successful, one can try to identify important parts of the image via heat map techniques or other interpretation techniques, but this analysis will be qualitative by nature and not necessarily interpretable by humans despite the best efforts (Samek et al. 2021). An alternative is to extract the coordinates of the important components of the image like defects, dislocations, etc. via computer vision techniques. Then one can use the same ML algorithms as used for MD, DDD and other particle based simulations or form various density measures with the equation discovery methods we presented in this paper.

Conclusion

Considering the successful applications of machine learning (ML) in recent years across many fields of science, we have given a short overview of the general process focusing on what should be taken into account when applying ML to mechanics/physics-related questions. We highlighted possible routes of applications in the field of dislocations dynamics models, atomistic models for plasticity in glasses and model discovery. In our opinion, it is vital to incorporate the underlying symmetry of the physical problem in the ML model in order to build a stable model with fewer training data. In order to advance our understanding of mechanics, ML should not be only pursued with the aim of increasing accuracy, but should also focus on the use of interpretable features and estimators. This combination of first principles and data science promises to be a highly rewarding endeavour to expand our knowledge and develop novel data driven materials theories.

Abbreviations

ML machine learning; DDD discrete dislocation dynamics; CDD continuum dislocation dynamics; CP crystal plasticity; PDE partial differential equations

Acknowledgements

We thank Michael Zaiser for useful discussions

Authors' contributions

The authors jointly wrote the manuscript and contributed equally. All authors read and approved the final manuscript.

Funding

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 377472739/GRK 2423/1-2019 and via the Grant no. 1 ZA 171/14-1. The authors are grateful for this support. Open Access funding enabled and organized by Projekt DEAL.

Availability of data and materials

Not applicable.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details

¹Institute of Materials Simulation, Department of Materials Science and Engineering, Friedrich-Alexander-University Erlangen-Nuremberg, Fürth, Germany. ²Center for Complexity and Biosystems, Department of Physics, University of Milan, Milan, Italy. ³Istituto di Chimica della Materia Condensata e di Tecnologie per l'Energia, CNR - Consiglio Nazionale delle Ricerche, Milan, Italy.

Received: 24 May 2021 Accepted: 24 June 2021

Published online: 01 September 2021

References

- V. Bapst, T. Keck, A. Grabska-Barwińska, C. Donner, E. D. Cubuk, S. S. Schoenholz, A. Obika, A. W. Nelson, T. Back, D. Hassabis, et al., Unveiling the predictive power of static structure in glassy systems. *Nat. Phys.* **16**(4), 448–54 (2020)
- J. Behler, M. Parrinello, Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**(14), 146401 (2007)
- N. Bertin, R. B. Sills, W. Cai, Frontiers in the simulation of dislocations. *Annu. Rev. Mater. Res.* **50**, 437–64 (2020)
- R. B. Bird, R. C. Armstrong, O. Hassager, Dynamics of polymeric liquids. *Fluid Mech.* **1**, 255–479 (1987)
- S. L. Brunton, J. L. Proctor, J. N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl. Acad. Sci.* **113**(15), 3932–3937 (2016)
- V. Bulatov, W. Cai, *Computer Simulations of Dislocations*, vol. 3. (Oxford University Press, Oxford, 2006). <https://doi.org/10.1093/oso/9780198526148.001.0001>
- K. Champion, B. Lusch, J. N. Kutz, S. L. Brunton, Data-driven discovery of coordinates and governing equations. *Proc. Natl. Acad. Sci.* **116**(45), 22445–22451 (2019)
- S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt, K.-R. Müller, Machine learning of accurate energy-conserving molecular force fields. *Sci Adv.* **3**(5), 1603015 (2017)
- E. D. Cubuk, S. S. Schoenholz, J. M. Rieser, B. D. Malone, J. Rottler, D. J. Durian, E. Kaxiras, A. J. Liu, Identifying structural flow defects in disordered solids using machine-learning methods. *Phys. Rev. Lett.* **114**(10), 108001 (2015)

- B. de Silva, K. Champion, M. Quade, J.-C. Loiseau, J. Kutz, S. Brunton, Pysindy: A python package for the sparse identification of nonlinear dynamical systems from data. *J. Open Source Softw.* **5**(49), 2104 (2020). <https://doi.org/10.21105/joss.02104>
- M. L. Falk, J. S. Langer, Dynamics of viscoplastic deformation in amorphous solids. *Phys. Rev. E.* **57**(6), 7192 (1998)
- Z. Fan, E. Ma, Predicting orientation-dependent plastic susceptibility from static structure in amorphous solids via deep learning. *Nat. Commun.* **12**(1), 1–13 (2021)
- H. Fan, Q. Wang, J. A. El-Awady, D. Raabe, M. Zaiser, Strain rate dependency of dislocation plasticity. *Nat. Commun.* **12**(1), 1–11 (2021)
- M. Gastegger, K. T. Schütt, K.-R. Müller, Machine learning of solvent effects on molecular spectra and reactions. *arXiv preprint arXiv:2010.14942* (2020)
- L. Landau, E. Lifshitz, *Theory of elasticity*. 1986. *Course Theor. Phys.*, 1–32 (1986)
- M. F. Langer, A. Goeßmann, M. Rupp, Representations of molecules and materials for interpolation of quantum-mechanical simulations via machine learning. *arXiv preprint arXiv:2003.12081* (2020)
- S. Lapuschkin, S. Wäldchen, A. Binder, G. Montavon, W. Samek, K.-R. Müller, Unmasking clever hans predictors and assessing what machines really learn. *Nat. Commun.* **10**(1), 1–8 (2019)
- C. W. Macosko, *Rheology principles*. *Meas. Appl.*, 109–174 (1994)
- M. C. Messner, M. Rhee, A. Arsenlis, N. R. Barton, A crystal plasticity model for slip in hexagonal close packed metals based on discrete dislocation simulations. *Model. Simul. Mater. Sci. Eng.* **25**(4), 044001 (2017)
- A. Pandey, R. Pokharel, Machine learning enabled surrogate crystal plasticity model for spatially resolved 3d orientation evolution under uniaxial tension. *arXiv preprint arXiv:2005.00951* (2020)
- S. Rudy, A. Alla, S. L. Brunton, J. N. Kutz, Data-driven identification of parametric partial differential equations. *SIAM J. Appl. Dyn. Syst.* **18**(2), 643–660 (2019)
- S. H. Rudy, S. L. Brunton, J. L. Proctor, J. N. Kutz, Data-driven discovery of partial differential equations. *Sci. Adv.* **3**(4), 1602614 (2017)
- M. Rupp, A. Tkatchenko, K.-R. Müller, O. A. Von Lilienfeld, Fast and accurate modeling of molecular atomization energies with machine learning. *Phys. Rev. Lett.* **108**(5), 058301 (2012)
- H. Salmenjoki, M. J. Alava, L. Laurson, Machine learning plastic deformation of crystals. *Nat. Commun.* **9**(1), 1–7 (2018)
- H. Salmenjoki, L. Laurson, M. J. Alava, Probing the transition from dislocation jamming to pinning by machine learning. *Mater. Theory.* **4**(1), 1–16 (2020)
- W. Samek, G. Montavon, S. Lapuschkin, C. J. Anders, K.-R. Müller, Explaining deep neural networks and beyond: A review of methods and applications. *Proc. IEEE.* **109**(3), 247–278 (2021)
- M. Sarvilahti, A. Skaugen, L. Laurson, Machine learning depinning of dislocation pileups. *APL Mater.* **8**(10), 101109 (2020)
- K. T. Schütt, P.-J. Kindermans, H. E. Saucedo, S. Chmiela, A. Tkatchenko, K.-R. Müller, Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. *arXiv preprint arXiv:1706.08566* (2017)
- D. Steinberger, H. Song, S. Sandfeld, Machine learning-based classification of dislocation microstructures. *Front. Mater.* **6**, 141 (2019)
- R. Tibshirani, Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B (Methodol.)* **58**(1), 267–288 (1996)
- O. T. Unke, S. Chmiela, H. E. Saucedo, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, K.-R. Müller, Machine learning force fields. *arXiv preprint arXiv:2010.07067* (2020)
- Q. Wang, J. Ding, L. Zhang, E. Podryabinkin, A. Shapeev, E. Ma, Predicting the propensity for thermally activated β events in metallic glasses via interpretable machine learning. *NPJ Comput. Mater.* **6**(1), 1–12 (2020)
- Q. Wang, A. Jain, A transferable machine-learning framework linking interstice distribution and plastic heterogeneity in metallic glasses. *Nat. Commun.* **10**(1), 1–11 (2019)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)