

https://doi.org/10.1038/s44334-024-00003-y

Data-driven modeling of processstructure-property relationships in metal additive manufacturing

Zhaoyang Hu & Wentao Yan 🖂

Check for updates

Metal additive manufacturing (AM) faces challenges in rapid selection and optimization of manufacturing parameters for desired part quality. As a more efficient alternative to experiments and high-fidelity physics-based models, data-driven modeling is effective in understanding process–structure–property relationships. This brief review explores data-driven modeling in metal AM, focusing on "process", "structure", and "property", further identifying limitations in current applications and accordingly presenting future outlook on the possible advancements in this domain.

Metal additive manufacturing (AM) has developed rapidly in recent decades due to its incomparable capabilities in fabricating parts. Specifically, instead of removing materials from a solid block in traditional subtractive methods, AM adds materials layer-by-layer to produce the desired geometry from a computer-aided-design (CAD) model. Such a precise layered approach allows for the creation of intricate structures with a high degree of customization and minimal material waste. However, the as-built parts by AM may not always possess satisfactory mechanical properties. This mostly stems from improper selection of process parameters, undesirable microstructure evolution, and unexpected defect formation, all of which essentially attribute to the incomplete understanding of process–structure–property (PSP) relationships in metal AM.

The challenge in establishing connections between "process", "structure", and "property" arises from various physical phenomena occurring simultaneously or sequentially, including but not limited to powder dynamics, laser ray reflection, heat transfer, fluid flow, phase transition between solid, liquid, and gas, solid-state phase transformation (SSPT), plastic deformation, fracture, and so on. These physical phenomena interact with each other, resulting in extremely complex PSP relationships. For example, under steep temperature gradients and frequent heating and cooling cycles, certain materials may undergo SSPT. Such a phenomenon could result in unusual thermal stress and local deformation, hence impacting mechanical properties like stiffness and strength, which may further lead to geometric inaccuracy and cracking. In addition, the gradients of temperature and the temperature history are determined by various factors, including laser power, scan speed, scan strategy, and material properties such as energy absorption, thermal conductivity, and specific heat.

A straightforward way to advance the understanding of PSP is to rationally design and carefully conduct experiments (particularly in situ ones) and/or develop high-fidelity physics-based simulations incorporating most if not all the major physical factors¹. However, experiments are costly while performing multi-physics simulations is also time-consuming. Such a dilemma hinders further investigations on PSP. More importantly, these two approaches are not feasible in the rapid selection and optimization of the manufacturing parameters, which is a critical requirement in industrial applications. In this case, leveraging data-driven models is believed to be a promising solution, especially in the era when machine learning techniques are well developed. Compared to traditional manufacturing processes like casting, AM significantly enhances design freedom and customization through its layer-by-layer manufacturing scheme. This manufacturing ability enables the production of complex geometries and internal structures that are beyond the capabilities of traditional methods. However, the layer-by-layer manufacturing scheme also introduces multiple challenges in data-driven modeling. Firstly, the complex combination of process parameters in AM allows for unprecedented flexibility and variability in manufacturing; however, this flexibility and variability also result in a large and high-dimensional data space that demands rigorous and thorough data collection for effective data-driven models. Secondly, quality inconsistencies inherent in AM, such as variations in porosity, surface roughness, and grain structures, challenge the reliability of collected data. Thirdly, the complex physics involved in AM processes makes it difficult to develop physics-informed data-driven models, hence challenging the physical interpretability of these models. Lastly, the developing but still nascent understanding of AM impedes the establishment of a unified experimental platform and simulation setup, which undermines the accuracy and reliability of collected data, further adding to the burden of data collection. Nevertheless, these challenges present opportunities for leveraging advanced data-driven techniques to fully explore the potential of AM. By refining data-driven modeling practices, these challenges can be mitigated to help enhance the AM for improved performance and reliability. Successful applications of data-driven modeling for AM have proven the appealing potential to optimize process parameters, to predict microstructure evolution, and to fine-tune the properties of final parts.

Recently, some review papers elaborate on the utilization of data-driven models in AM. Wang et al.² reviewed the machine learning applications in

Department of Mechanical Engineering, National University of Singapore, Singapore, 117575, Singapore. Me-mail: mpeyanw@nus.edu.sg

structural design for AM parts, process parameter optimization, process monitoring, and quality control. In the review by Jin et al.³, the authors focused on the machine learning-assisted topology optimization for structural design and the potential of machine learning in material design for AM, as well as machine learning applications in process parameter optimization and in situ anomaly detection. Meng et al.⁴ reviewed process parameters optimization, defect detection, quality and property prediction, and geometric deviations control with the aid of machine learning. Qi et al.⁵ specifically focused on the applications of neural networks in AM, e.g., structural design and geometric compensation, in situ monitoring, and connecting process parameters with properties of final parts. Wang et al.⁶ reviewed machine learning-based process modeling, in situ defect detection, and process optimization and control. Liu et al.7 emphasized the machine learning techniques for process and performance optimization. Li et al.8 presented the machine learning applications in the pre-processing design stage, e.g., structural design, during the processing stage, e.g., process parameter optimization and in-process monitoring optimization, and in the post-processing stage, e.g., surface quality prediction. Wang et al.9 reviewed the data-driven process modeling, e.g., geometry of molten pool and bead, data-driven structure modeling, e.g., grain structure and geometric distortion, and data-driven mechanical property modeling, e.g., tensile strength. Particularly, they gave a detailed introduction of various data-driven models being used and discussed how to select suitable data-driven models for PSP modeling. Kouraytem et al.¹⁰ reviewed data-driven applications for



Fig. 1 | Schematic of data-driven modeling of PSP in metal AM. Data-driven models are leveraged to help comprehensively understand PSP and rapidly establish mappings for PSP under the interaction of complex physical phenomena occurring concurrently or sequentially. Figures are reprinted from refs. 31,33,69–71 with permission from Springer Nature and Elsevier.

process-structure relationship, structure-property relationship, and the vision for data-driven PSP adopted from Yan et al.¹¹ and Wang et al.¹².

Different from the existing reviews in the literature, this paper is aimed not only at presenting representative successful data-driven applications in studying PSP, but also at elaborating on the current challenges of linking PSP through data-driven techniques and discussing the future prospects and possibilities of data-driven modeling in AM. Instead of a very comprehensive overview like that by Wang et al.⁹, this review is confined to classical and pioneering literatures and provides a more concise yet insightful exploration of landmark studies in the field. Most data-driven techniques are broadly applicable to various AM techniques. For example, a data-driven application developed for one specific AM technique, e.g., directed energy deposition (DED), can often be generalized to another AM technique, e.g., laser powder bed fusion (LPBF). Therefore, we avoid confining our scope to data-driven applications of a specific AM technique. It is also important to highlight that this cross-applicability not only showcases the generalization capacity of data-driven applications in AM but also fosters advancements in comprehending PSP as a whole. In the following, successful data-driven applications in the literature are firstly reviewed from process modeling, e.g., process parameter optimization and characteristics of molten pool, to structure modeling, e.g., evolution of grain structure, and property modeling, e.g., prediction of ultimate tensile strength (UTS) and yield strength, as schematically illustrated in Fig. 1. Particularly, we try to leverage these examples to shed light on the essential reasons behind the indispensability of data-driven modeling for PSP. Furthermore, we try to specifically identify the limitations of current data-driven applications and outlook future directions to address these issues.

Data-driven applications for process, structure, and property modeling

In this section, we discuss representative data-driven modeling for "process", "structure", and "property". Figure 2 serves as a preview, illustrating specific applications related to these aspects.

Data-driven modeling of process

To model AM processes, researchers commonly use thermal-fluid simulations within the computational fluid dynamics (CFD) framework. This approach offers two key advantages: (1) it eliminates the need for numerous trial-and-error experiments, thus largely reducing costs; (2) it is more accurate than conventional pure thermal simulations based on the finite element method (FEM) due to its incorporation of molten pool flow dynamics and other significant physical factors¹³. However, it is noteworthy that such high-fidelity simulations are inherently time-consuming, e.g., the computational time of a typical simulation case of laser melting a single track is on the order of hundreds of CPU hours. Consequently, researchers have



Fig. 2 | **Schematic of data-driven applications for process, structure, and property modeling. a** Data-driven process modeling with training data mainly from simulations (in red). **b** Data-driven structure modeling with training data mainly from simulations (in red). **c** Data-driven property modeling (from "process" to "property") with training data mainly from experiments (in red). **d** Data-driven property modeling (from "structure" to "property") with training data mainly from experiments (in red).

leveraged data-driven techniques to directly establish relationships between process parameters and corresponding process features, such as the geometry of the molten pool, as illustrated in Fig. 2a. In this section, our focus narrows to a handful of seminal studies out of the extensive literature, so that we offer a glimpse into representative applications for data-driven modeling of process.

The geometry of the molten pool is commonly adopted as an indicator of the quality of as-built parts, as concerns like porosity and lack-of-fusion are closely associated with the characteristics of the molten pool. Therefore, capturing the relationships between the process parameters and geometry of the molten pool can help optimize the process parameters for better quality of as-built parts. In addition, since the geometry of the molten pool is easy to measure in both experiments and high-fidelity simulations, this quantity becomes a widely used metric in metal AM. With limited experimental data and high-fidelity simulations available, Tapia et al.¹⁴ built a Gaussian process-based surrogate model to predict the molten pool depth given different combinations of laser power, scan speed, and laser beam size in LPBF. The proposed surrogate model was then used in AM processes planning to obtain the desirable conduction mode rather than the keyhole mode for the molten pool. The same methodology was also used in Tapia et al.¹⁵ to predict the porosity and to find an optimized combination of laser power and scan speed to achieve low porosity in LPBF. It should be noted that Gaussian process regression is widely used in such prediction tasks, as it is a non-parametric statistical tool capable of accurately capturing nonlinear mappings from inputs to outputs¹⁶, without demanding large amounts of training data. Xie et al.¹⁷ applied a Gaussian process regression model to predict molten pool and keyhole sizes, and a deep neural network (DNN) to recognize molten pool melting regime, keyhole stability, and keyhole type, where the database comes from a high-fidelity thermal-fluid flow model¹⁸. For selective laser sintering, Garg et al.¹⁹ proposed a new ensemble-based multi-gene genetic programming model that uses statistical approaches of stepwise regression and classification methods, including support vector machines, Bayesian classifiers, and DNNs, for improving its generalization. They used experimental data to train the model and predicted the open porosity based on three uncertain process parameters: layer thickness, laser power, and scan speed. Garg et al.²⁰ also applied the same methodology to predict the variation of bead width with respect to layer thickness, laser power, and scan speed for LPBF. For electron beam powder bed fusion, Aoyagi et al.²¹ utilized the support vector machine to predict the porosity and surface roughness from the combinations of beam current and scan speed. Based on the experimental data obtained from literature, Akbari et al.²² used various machine learning algorithms, i.e., neural network, gradient boosting, random forest, Gaussian process model, support vector machine, and ridge/lasso/logistic regression to predict the geometry of the molten pool, and also for classification tasks that identify the defect modes within the molten pool, such as lack of fusion and keyhole porosity. The authors concluded that neural network, gradient boosting, and random forest outperform the other machine learning algorithms on the two tasks. Some researchers also utilized data-driven techniques to predict the bead geometry. Xiong et al.²³ investigated the relationship between the process parameters and the bead geometry through a neural network and a secondorder regression for gas metal arc welding. While these data-driven models demonstrate good accuracy in prediction under carefully controlled lab/ simulation conditions, their applicability and transferability in real manufacturing situations remain questionable. The practical viability of these approaches hinges on several factors, including the robustness of datadriven models across different material types, machine precision, and environmental conditions in actual manufacturing settings. Assessing the reliability and adaptability of these approaches outside controlled conditions, as discussed in a later section ("Complete integration of PSP in datadriven modeling") through uncertainty quantification, is an essential step to validate their practical utility.

Besides, thermal history during AM processes significantly influences microstructure evolution and mechanical properties by inducing microscopic phase transformations and grain growth, along with the evolution of thermal and residual stresses during the laver-by-laver deposition. However, it is very challenging, if not impossible, to experimentally measure the threedimensional (3D) temperature field. Therefore, accurate prediction of thermal history is the foundation to optimize microstructural characteristics and mechanical properties. Mozaffar et al.²⁴ presented a stacked recurrent neural network with a gated recurrent unit formulation to predict the thermal history of any given point during the DED process. Roy and Wodo²⁵ developed a neutral network-based surrogate model for predicting the thermal history during fused filament fabrication process. The datasets in refs. 24,25 are both generated through physics-based simulations using the FEM. However, FEM-based thermal analysis incorporates only heat conduction and normally omits the effects of fluid flow, powder melting, etc. Zhu et al.²⁶ proposed a physics-informed neural network that makes use of both training data from simulations and physical principles, i.e., conservation laws of momentum, mass, and energy and corresponding boundary conditions, to predict temperature field and molten pool dynamics. Although the authors did consider complex thermal behaviors beyond heat conduction, the free surface of the molten pool was not captured, which compromised the accuracy. The aforementioned oversimplifications in refs. 24-26 lead to somewhat inaccurate temperature field, and ref. 13 presented detailed assessment of inaccuracies in FEM-based heat transfer models by benchmarking various models against experimental results and theoretically analyzing the assumptions of each model. Consequently, some researchers chose to develop data-driven models on the basis of high-fidelity CFD models. Chen et al.²⁷ used a number of isotherms instead of a patternbased field to represent the temperature profile around the molten pool, and respectively applied Gaussian process regression, quadratic regression, linear regression, and support vector regression to predict the isotherm dimensions based on different combinations of absorbed laser power and scan speed. Only a small amount of training data from high-fidelity thermalfluid CFD model is needed. The predicted isotherms are used to reconstruct the temperature field, which is then applied in the thermal stress and grain growth models. Through verification, Gaussian process regression and quadratic regression perform better than the other two algorithms. Strayer et al.²⁸ presented a CFD imposed FEM framework to simulate the thermal profile in LPBF. The thermal profile near the molten pool is calculated through the CFD simulations, and heat diffusion elsewhere is computed via the FEM. To further accelerate the simulation, DNNs are trained with the CFD simulations as the ground truth, and the trained DNNs replace the CFD simulations to yield the thermal profile near the molten pool. Hemmasian et al.²⁹ trained a 3D convolutional neural network (CNN) via the CFD simulations for single track LPBF. This CNN outputs the thermal profile based on the inputs of laser power, scan speed, and time step. Since the incorporation of time step as one of the inputs, the thermal profile for any arbitrary time step can be obtained immediately without iterations.

Data-driven modeling of structure

Given the scarcity of in situ experimental results, physics-based simulations stand out as the primary avenue for studying microstructure evolution in metal AM. Physics-based models commonly use kinetic Monte Carlo method³⁰, cellular automata method^{31,32}, or phase-field method³³. However, due to sufficiently small mesh size needed to resolve the microstructure features, physics-based simulations are computationally expensive. In this case, data-driven modeling can be utilized to accelerate the computation of microstructure evolution so as to more easily correlate process parameters with the formation and characteristics of grain structures, as illustrated in Fig. 2b. Despite the extensive applications for data-driven process modeling discussed in the section "Data-driven modeling of process", there are limited literatures available on data-driven structure modeling. Popova et al.³⁴ proposed a low-dimensional, data-driven approach to link manufacturing processes with predicted microstructure. Specifically, the authors applied the multivariate polynomial regression to establish the relationship between various parameters, such as temperature field, scan speed, and dimensions of the molten pool and heat-affected zone, along with the chord length distribution of the grains calculated using the Potts-kinetic Monte Carlo

method³⁵. By conducting the principal component analysis on the chord length distribution, they found that only four principal components are already sufficient to accurately reconstruct the chord length distribution. Therefore, the regression task shifts from predicting the detailed chord length distribution to simply predicting the four principal components of the distribution. This simpler, low-dimensional regression reduces both computational time and storage demands. Xue et al.³⁶ leveraged a physicsembedded graph neural network (PEGN) to accelerate the phase-field simulations for grain growth³³ in LPBF, that is, mapping the phase-field physics into the framework of the graph neural network through incorporating each individual grain and adjacent grain interactions. This incorporation is achieved by interpreting the polycrystalline structure as an undirected graph, with associated node features such as the volume or the centroid position of grains and edge features such as contact faces and thermal conductivity between grains. The same PEGN is also applied to solve temperature field and liquid/solid phase fraction, and this proposed framework is confirmed to be at least 50 times faster than the direct numerical simulation in computation. Utilizing small-scale phase-field simulations, Choi et al.³⁷ trained a 3D CNN to predict the grain orientation of the next time step given the current grain orientation and the temperature fields at the current and next time step. To perform full simulations using the trained small-scale CNN, the local grain structure at the current laser position is sampled from the entire domain, and only the local temperature field within this small volume of interest needs to be calculated to enable the CNN to generate the resulting grain orientations. Through spatio-temporal composition, integrating small-scale predictions with the movement of the laser, the grain orientation of the entire domain is updated progressively.

Data-driven modeling of property

As the final step of PSP, "property" is affected by both "process" and "structure". Particularly, we would like to clarify that besides strength, fatigue resistance and others, thermal/residual stresses are also considered as "property" here, while some researchers classify them as "structure". The tensile strength is not only determined by the microstructure, but also influenced by the thermal/residual stresses originating mainly from the large temperature gradient during the process. Therefore, performing full-scale simulations together with comprehensive experimental observations is indispensable to provide insight into the impact of process parameters and microstructure on mechanical properties. However, it is extremely challenging to perform fast and accurate full-scale simulations with current computational resources as well as carrying out massive experiments. In such circumstances, leveraging data-driven techniques provides valuable assistance in overcoming these difficulties. In this section, we review some representative data-driven applications for property modeling. Figure 2c illustrates that many data-driven models establish a direct relationship between "process" and "property", often overlooking "structure". In contrast, as illustrated in Fig. 2d, a minority of models exclusively use "structure" information to predict "property", theoretically offering a more rigorous approach as "property" is directly determined by "structure" and indirectly by "process". Despite the theoretical rigorousness of the PSP framework, current research predominantly focuses on direct mappings from "process" (e.g., laser power and scan speed) to "property" (e.g., ultimate tensile strength, yield strength). This direct approach achieves good prediction accuracy but typically bypasses detailed "structure" information. Challenges in obtaining comprehensive microstructural data, such as grain morphology, through experimental techniques like two-dimensional (2D) electron backscattered diffraction (EBSD), hinder accurate and in-depth analysis, especially for materials with non-equiaxed grain shapes common in AM processes³⁸. Although 3D EBSD offers more accurate morphology insights, its extensive requirements in labor, cost, and time limit its widespread use. Similarly, in simulations, 2D physics-based models inadequately capture grain growth during AM processes, while 3D models are extremely timeconsuming³³, even for single-track simulations. Given these challenges, establishing direct process-property linkage is understandable and justifiable. Additionally, this approach is more of practical value in engineering applications, as it provides direct guidance on process parameter selection to obtain desired properties.

Tensile properties such as UTS and yield strength of AM parts are critical mechanical properties with significant implications for the structural integrity and performance during engineering service. Accurate prediction of such properties is essential in assessing the material's ability to withstand maximum service loading before failure. Zhang et al.³⁹ proposed a concatenated neural network, including a long short-term memory (LSTM) network and a fully connected neural network, to predict the UTS based on 144 specimens by fused deposition modeling (FDM). In the LSTM network, each LSTM cell corresponds to the experimentally measured temperature and vibration data of a FDM printing layer, and the LSTM outputs feed into the fully-connected network, along with process parameters and material properties for final UTS prediction. Although the material is polylactic acid (PLA) not metals, the concatenated neural network is worth mentioning, as it reflects the inter-layer influence during printing by communications between LSTM cells. Marmarelis and Ghanem⁴⁰ utilized diffusion maps on a lower-dimensional Riemannian manifold to infer the dependence between process parameters (hatch spacing, scan speed, and laser power) and part properties (UTS, yield strength, and elongation) based on 51 AlSi10Mg specimens by LPBF. On the lower manifold the authors further conducted stochastic optimization of the process parameters to achieve minimum chance of failure. Ponticelli et al.⁴¹ proposed a fuzzy-multi-objective genetic model using the NSGA-II algorithm⁴². They used the model to find the best combination of building orientation and volumetric energy density, with the latter being directly influenced by laser power, scan speed, hatch spacing, and layer thickness, for the highest mechanical performances, i.e., UTS, roughness, and hardness, based on 30 AlSi10Mg specimens by LPBF. Hertlein et al.43 developed a Bayesian neural network to relate process parameters (laser power, scan speed, hatch spacing, and layer thickness) and part properties (density, hardness, top layer surface roughness, UTS in the build direction and perpendicular to the build direction) based on 349 316L stainless steel specimens by LPBF collected from literature. Unlike traditional neural networks that offer single value estimates, Bayesian neural networks provide predictions as probability distributions, which indicate both the most likely value and the uncertainty level for each prediction. This effectively enhances the capability for uncertainty quantification of the predictions. Xie et al.44 predicted the distribution of UTS, yield strength, and elongation from the local thermal history via the infrared (IR) thermal measurement by integrating wavelet transforms and CNNs based on 135 selected local regions from 12 Inconel 718 specimens built by DED. Similarly, Fang et al.45 used CNNs to predict UTS, yield strength, failure stress, and Young's modulus of Inconel 718 by DED directly from the thermal history obtained from the FEM simulation rather than the IR measurement.

Fatigue life of AM parts directly influences the long-term structural integrity and durability. Employing data-driven modeling enables rapidly assessing the fatigue performance on different selection of process parameters. Zhang et al.⁴⁶ adopted a neuro-fuzzy-based neural network to realize high-cycle fatigue life prediction for 316L stainless steel by LPBF. Both process-based model and property-based model are developed for fatigue life prediction. The inputs of the process-based model consist of cyclic stresses, process parameters (laser power, scan speed, and layer thickness), and post-processing treatments (annealing and hot isostatic pressing) while the inputs of the property-based model consist of UTS and elongation. Zhan and Li47 first derived a continuum damage mechanics-based formulation for fatigue life calculation and then implemented three machine learning techniques, i.e., neural network, random forest, and support vector machine, to predict the fatigue life of 316L stainless steel by LPBF with sufficient data generated by the theoretical formulation. In this study, the inputs for fatigue life prediction include four process parameters (laser power, scan speed, hatch spacing, and layer thickness) and two fatigue loading parameters (maximum stress and stress ratio). These efforts on predicting the fatigue life directly from "process" while neglecting "structure" further implies practical complexities and challenges in establishing complete PSP for property

prediction. The scarcity of reliable physics-based models to capture fatigue behaviors based on microstructure adds to these challenges. Even when comprehensive microstructural data can be obtained, the prediction of fatigue life remains less reliable. Therefore, combined efforts to integrate "structure" and develop advanced physics-based models for fatigue are necessary to achieve better prediction performance.

Some researchers develop data-driven models to predict mechanical properties based on microstructure instead of process parameters. By applying random forest and watershed segmentation, Kusano et al.48 first analyzed the morphology and size of the microstructure of Ti-6Al-4V alloy produced by LPBF followed by post-heat treatment, using scanning electron microscopy images. They then used multiple linear regression analysis to establish a connection between the microstructure and both UTS and yield strength. Based on the microstructure dataset generated by the elastoviscoplastic fast Fourier transform approach in ref. 49, Herriott and Spear⁵⁰ implemented ridge regression, XGBoost, and CNN to predict yield strength of 316L stainless steel based on microstructure information. The inputs for ridge regression and XGBoost consist of morphological and crystallographic features to characterize the local microstructure. In contrast, CNN utilizes 3D images of the microstructure as the inputs, with crystal orientation assigned to each voxel as the primary feature, complemented by taking the micromechanical Taylor factor as the auxiliary feature. In this setup, CNN outperforms the other two methods. This superior performance may be attributed to the remarkable ability of CNN to extract spatial information from images.

Despite these advancements, several issues are worth mentioning. The indices of mechanical properties in current data-driven models remain limited, with very few reports on properties such as creep, high-temperature properties, and high-strain-rate properties, not to mention functional properties like electrical conductivity. However, these properties are particularly important in target applications such as aerospace engineering. Moreover, the amount of training data, whether from experiments or simulations, is generally small. This scarcity is due to the high cost of systematic material characterizations and tests. These challenges will be further discussed in a later section ("Deficiency and reliability of training data").

Research limitations and future outlook

From the above review, data-driven models leverage datasets obtained from simulations and/or experiments to identify complex correlations between "process" and "structure", between "process" and "property", and between "structure" and "property". These successful applications foster a deeper comprehension of PSP and facilitates the optimization of process parameters. This would further enable researchers to tailor AM materials and design AM parts with enhanced engineering performance, such as improved UTS and fatigue life. In addition, a comprehensive understanding of the relationships between process parameters, microstructure, and mechanical properties aids in quality control, ensuring the quality consistency of AM parts. Such insight is significant for industries that demand precision and reliability, such as aerospace and healthcare. However, there are some limitations in current data-driven applications that may impede further understanding of PSP, and we elaborate on these limitations and outlook possible solutions worthy of endeavor in the following content, as illustrated in Fig. 3.

Complete integration of PSP in data-driven modeling

As mentioned in the section "Data-driven modeling of property", most of the reported data-driven applications for property modeling^{39–41,43–47} tend to predict mechanical properties directly from process parameters, and the other applications^{48,50} map the relationships between microstructure and mechanical properties. Nevertheless, since mechanical properties are affected by a multitude of factors, including microstructure evolution and temperature profile, integrating both process data and microstructure information tends to yield a more comprehensive understanding and consequently better prediction of mechanical properties. In this case, PSP is



Fig. 3 | Future outlook for advancing data-driven PSP, including complete linkage of "process", "structure", and "property" with interpretation and uncertainty quantification, data standardization and augmentation, and inverse design applications.

fully leveraged, and this holistic approach acknowledges the interplay of various factors affecting mechanical properties, enhancing accuracy and reliability in predicting mechanical properties.

Particularly, thermal/residual stresses arising from rapid heating and cooling cycles during AM processes profoundly affect mechanical properties, such as UTS, yield strength, ductility, and fatigue life. This is actually one of the most concerned problems in current industrial applications. Firstly, high-level thermal/residual stresses tend to concentrate at specific locations, such as defects and geometric irregularities. Such stress concentrations may lead to crack initiation and propagation, significantly compromising the overall mechanical integrity. Secondly, the constant presence of high thermal/residual stresses often results in macroscopic dimensional distortions of as-built parts⁵¹. Substantial geometric deviations from the intended geometry adversely impact the desired mechanical properties and overall mechanical performance. Thirdly, complex interactions occur between microstructure evolution and microscopic residual stresses. For example, unusual microstructure formation during nonequilibrium solidification and solid-state phase transformations affect the distribution and magnitude of residual stresses⁵², and thermal/residual stresses can also induce microstructure evolution, e.g., recrystallization. These further complicate the relationships between microstructure, residual stresses, and mechanical properties⁵³. Therefore, comprehensively considering thermal/residual stresses is crucial for advancing the understanding of PSP to the next level, yet there is a notable absence of relevant data-driven models for multi-scale stress analysis, which is a critical gap in current research efforts.

Besides, another critical issue in AM is quality inconsistency of as-built parts, which significantly hampers industrial applications. The inconsistency initially arises from uncertainties of process parameters and material properties, for example, due to machine calibration and powder oxidation⁵⁴. Such uncertainties cannot be eliminated at the very beginning and would naturally propagate and accumulate along PSP, and ultimately affect the final mechanical properties^{55,56}. As a result, there is a critical demand for collecting reliable and consistent datasets involving "process", "structure", and "property" and carefully conducting uncertainty quantification on data-driven PSP. Both uncertainties from datasets and prediction variances from data-driven models should be assessed. This can systematically reveal the impacts of uncertain factors, e.g., laser power and thermal conductivity, on the outputs of interest, e.g., molten pool size and UTS. Accordingly, the understanding of PSP will be further improved, resolving conflicting experimental observations and improving the prediction accuracy of both the physics-based and data-driven models. More importantly, valuable guidance can be provided to experiments and machine development to better ensure quality consistency. For example, some significant factors, like laser diameter, should be more carefully calibrated and precisely controlled if possible for each printing job, while the efforts and investment on insignificant factors can be reduced.

Active learning is another data-driven technique that can be leveraged effectively for linking PSP. In supervised learning, active learning reduces the need for a large number of labeled data pairs of inputs and outputs by selectively querying for labels on specific data pairs⁴. This selective inquiry enhances training efficiency and performance. For example, in investigating the relationship between process parameters, e.g., laser power and scan speed, and defect modes within the molten pool, e.g., keyhole porosity and lack of fusion, active learning can identify critical data points (only a portion of the whole dataset) that require labeling for defect modes during the training stage. This active learning-assisted procedure reduces the demand for a large labeled dataset in metal AM.

In addition, with superior capabilities in information comprehension, particularly by means of the advanced self-attention mechanism and parallel processing capabilities for efficient identification of internal relationships, large language models (LLMs) hold significant potential in establishing PSP for AM. Moreover, the interactive capabilities of LLMs through text-based or voice-based⁵⁷ prompts facilitate the seamless integration of the entire workflow-from design through manufacturing to performance evaluation⁵⁸. This capability would significantly improve the efficiency and decision-making processes in AM. Furthermore, with the active LLM community, fine-tuning open-source models such as LLaMa-13B⁵⁹ significantly reduces the effort required to customize more specialized models for specific AM tasks. These customized models are expected to surpass conventional LLMs, which are often inclined to simple analyses on the problems at hand rather than sophisticated investigations⁵⁸. Such oversimplification, for example in evaluating mechanical properties of as-built engineering structures, could potentially raise issues related to structure integrity.

Deficiency and reliability of training data

Data-driven modeling of PSP demands ample and reliable training data. However, the harsh conditions and complex physical mechanisms during AM processes collectively lead to the difficulty in obtaining comprehensive, reliable, and accurate data from experiments and simulations. Firstly, extreme conditions, e.g., high temperature and rapid solidification, highly affect the accuracy and reliability of data collection in experiments. In addition, conducting high-fidelity simulations of AM processes is extremely computationally demanding, necessitating substantial computational resources and storage. This is attributed to the integration of various complex physical mechanisms, including powder dynamics, laser ray reflection, heat transfer, and fluid flow, etc. The computational expense and the effectiveness of theoretical models for these physical mechanisms impede the acquisition of available data in the simulations. Moreover, due to machine calibration and other environmental factors, even adopting exactly the same process parameters but on different machines yields different experimental results⁶⁰. This inconsistency also applies to simulations as different physics-based models have different simplifications with various implementation details⁶⁰, which further highlights the significance of uncertainty quantification⁵⁵ mentioned in the section "Complete integration of PSP in data-driven modeling". Under such lack of available and reliable data derived from experiments and high-fidelity simulations, it is hard to compare the trustworthiness and utility of different data-driven models, hindering the selection of the most suitable one for linking PSP, as the scarcity of reliable data significantly undermines the model ability to capture the complexity of underlying PSP.

One effective approach to address the challenge of limited and unreliable data is through data sharing and standardization practices. By fostering a collaborative environment where researchers openly share their experimental results and simulation outcomes, the collective dataset grows more comprehensive, spanning sufficient variability in parameter space for metal AM. Furthermore, implementing standardized formats for data representation ensures compatibility, facilitating seamless integration of data acquired from different sources. This concerted effort in data sharing and standardization not only enhances the reliability of data-driven models but also promotes a more robust understanding of PSP across the scientific community. Developing more precise measurements, proposing more accurate and efficient theoretical models, and establishing an engaged community committed to data sharing and standardization are the fundamental steps to push forward the understanding of PSP to the next stage, and, in turn, optimizing the performance and reliability of AM techniques. Moreover, collecting FAIR (findable, accessible, interoperable, and reusable) data^{61,62} and establishing ontology-based knowledge map⁶³⁻⁶⁵ play a significant role in this endeavor. By structuring and integrating FAIR data within a formalized ontology, it becomes feasible to systematically map the interconnected relationships between process parameters, microstructure evolution, and resultant mechanical properties. This not only facilitates effective data sharing and standardization but also enhances automated data retrieval and reasoning⁶⁴, further improving the accuracy and robustness of data-driven PSP models. Interested readers are referred to the Additive Manufacturing Materials Database (AMMD⁶⁶) at National Institute of Standards and Technology, which is one such initiative aimed at providing a centralized data repository for the AM community.

Interpretation of data-driven-based PSP

While various neural networks have been developed to effectively investigate PSP, the "black-box" nature of neural networks, characterized by the lack of transparency in the mapping from input to output, largely hinders the interpretation and trustworthiness of the newly discovered PSP. To address this challenge, various strategies can be employed. Firstly, the implementation of feature attribution techniques offers a solution by identifying the specific input features that wield the largest influence on the corresponding output. This approach provides a clearer understanding of the influential factors, hence facilitating the understanding of PSP with enhanced interpretability. Secondly, the development of physics-informed neural networks (PINNs), via the integration of fundamental physical principles into neural networks, emerges as a promising direction^{26,36}. This not only improves the interpretability of the neural networks but also has the potential to boost the accuracy by aligning the learned relationships more closely with the governing physical principles. Further, with the ability to incorporate underlying physical principles (or governing equations) into the learning process, PINNs reduce the need for a large amount of training data, which, to some extent, could ease the challenge of limited and unreliable training data discussed in the section "Deficiency and reliability of training data". Nevertheless, due to the iterative nature of gradient descent and the computation of higher-order derivatives, especially when involving complex physical principles in PINNs, high computational cost and numerical instability, e.g., vanishing or exploding gradients, may arise. Hence, additional endeavors to enhance the performance of PINNs are required.

PSP applications in inverse design

Unlocking the potential for inverse design in metal AM highly depends on the understanding of PSP and powerful software tools. This is particularly significant for structural design, topology optimization, and the exploration of new material designs. Researchers can achieve desired outcomes by untangling the complex interplay among process parameters, microstructure, and mechanical properties. This advanced knowledge derived from PSP serves as a guide to achieve specific mechanical performance, through tailoring of process parameters, fine-tuning material compositions, and designing structural and topological configurations.

Despite the seamless connection between CAD and manufacturing in AM, challenges arise with the complex geometries generated through structural design and topology optimization. For example, printability issues, such as overhang and connectivity constraints, become prominent concerns⁶⁷. To address this, future studies should incorporate manufacturing feasibility into the optimization process, ensuring that the generated geometries not only exhibit structural efficiency but are also conducive to successful manufacturing.

Furthermore, the pursuit of new material designs introduces challenges due to the extremely large search domain. Efficient machine learning algorithms and substantial computational resources are essential to tackle these challenges. A recent study by Google⁶⁸ showcased the use of large-scale graph neural networks, achieving unprecedented efficiency in material discovery and surpassing human chemical intuition significantly. However, effective data-driven frameworks for material design remain notably scarce in the current literature. Therefore, there is a need to develop efficient machine learning algorithms that explore the design space while simultaneously making large-scale computations more accessible, possibly through cloud computing or distributed computing systems. This joint effort ensures the optimal utilization of advanced machine learning algorithms, contributing to a significant leap forward in data-driven material design.

Received: 21 February 2024; Accepted: 19 August 2024; Published online: 12 September 2024

References

- Wang, L., Guo, Q., Chen, L. & Yan, W. In-situ experimental and highfidelity modeling tools to advance understanding of metal additive manufacturing. *Int. J. Mach. Tools Manuf.* **193**, 104077 (2023).
- Wang, C., Tan, X. P., Tor, S. B. & Lim, C. S. Machine learning in additive manufacturing: state-of-the-art and perspectives. *Addit. Manuf.* 36, 101538 (2020).
- Jin, Z., Zhang, Z., Demir, K. & Gu, G. X. Machine learning for advanced additive manufacturing. *Matter* 3, 1541–1556 (2020).
- Meng, L. et al. Machine learning in additive manufacturing: a review. JOM 72, 2363–2377 (2020).
- Qi, X., Chen, G., Li, Y., Cheng, X. & Li, C. Applying neural-networkbased machine learning to additive manufacturing: Current applications, challenges, and future perspectives. *Engineering* 5, 721–729 (2019).
- Wang, P., Yang, Y. & Moghaddam, N. S. Process modeling in laser powder bed fusion towards defect detection and quality control via machine learning: The state-of-the-art and research challenges. *J. Manuf. Process.* **73**, 961–984 (2022).
- Liu, J. et al. A review of machine learning techniques for process and performance optimization in laser beam powder bed fusion additive manufacturing. *J. Intell. Manuf.* 34, 3249–3275 (2023).
- 8. Li, K. et al. A review of the multi-dimensional application of machine learning to improve the integrated intelligence of laser powder bed fusion. *J. Mater. Process. Technol.* **318**, 118032 (2023).
- Wang, Z. et al. Data-driven modeling of process, structure and property in additive manufacturing: A review and future directions. *J. Manuf. Process.* 77, 13–31 (2022).
- Kouraytem, N., Li, X., Tan, W., Kappes, B. & Spear, A. D. Modeling process-structure-property relationships in metal additive manufacturing: a review on physics-driven versus data-driven approaches. J. Phys.: Mater. 4, 032002 (2021).
- Yan, W. et al. Data-driven multi-scale multi-physics models to derive process-structure-property relationships for additive manufacturing. *Comput. Mech.* 61, 521–541 (2018).
- Wang, Z. et al. Uncertainty quantification in metallic additive manufacturing through physics-informed data-driven modeling. *JOM* 71, 2625–2634 (2019).
- Yan, W. et al. Data-driven characterization of thermal models for powder-bed-fusion additive manufacturing. *Addit. Manuf.* 36, 101503 (2020).
- Tapia, G., Khairallah, S., Matthews, M., King, W. E. & Elwany, A. Gaussian process-based surrogate modeling framework for process planning in laser powder-bed fusion additive manufacturing of 316L stainless steel. *Int. J. Adv. Manuf. Technol.* **94**, 3591–3603 (2018).
- Tapia, G., Elwany, A. H. & Sang, H. Prediction of porosity in metalbased additive manufacturing using spatial Gaussian process models. *Addit. Manuf.* **12**, 282–290 (2016).
- 16. Seeger, M. Gaussian processes for machine learning. *Int. J. Neural Syst.* **14**, 69–106 (2004).

- Xie, Z., Chen, F., Wang, L., Ge, W. & Yan, W. Data-driven prediction of keyhole features in metal additive manufacturing based on physicsbased simulation. *J. Intell. Manuf.* **35**, 2313–2326 (2024).
- Yan, W. et al. Multi-physics modeling of single/multiple-track defect mechanisms in electron beam selective melting. *Acta Mater.* 134, 324–333 (2017).
- Garg, A., Lam, J. S. L. & Savalani, M. M. A new computational intelligence approach in formulation of functional relationship of open porosity of the additive manufacturing process. *Int. J. Adv. Manuf. Technol.* **80**, 555–565 (2015).
- Garg, A., Tai, K. & Savalani, M. M. Formulation of bead width model of an SLM prototype using modified multi-gene genetic programming approach. *Int. J. Adv. Manuf. Technol.* **73**, 375–388 (2014).
- Aoyagi, K., Wang, H., Sudo, H. & Chiba, A. Simple method to construct process maps for additive manufacturing using a support vector machine. *Addit. Manuf.* 27, 353–362 (2019).
- Akbari, P. et al. MeltpoolNet: melt pool characteristic prediction in Metal Additive Manufacturing using machine learning. *Addit. Manuf.* 55, 102817 (2022).
- Xiong, J., Zhang, G., Hu, J. & Wu, L. Bead geometry prediction for robotic GMAW-based rapid manufacturing through a neural network and a second-order regression analysis. *J. Intell. Manuf.* 25, 157–163 (2014).
- Mozaffar, M. et al. Data-driven prediction of the high-dimensional thermal history in directed energy deposition processes via recurrent neural networks. *Manuf. Lett.* 18, 35–39 (2018).
- 25. Roy, M. & Wodo, O. Data-driven modeling of thermal history in additive manufacturing. *Addit. Manuf.* **32**, 101017 (2020).
- Zhu, Q., Liu, Z. & Yan, J. Machine learning for metal additive manufacturing: predicting temperature and melt pool fluid dynamics using physics-informed neural networks. *Comput. Mech.* 67, 619–635 (2021).
- Chen, F., Yang, M. & Yan, W. Data-driven prognostic model for temperature field in additive manufacturing based on the high-fidelity thermal-fluid flow simulation. *Comput. Methods Appl. Mech. Eng.* 392, 114652 (2022).
- Strayer, S. T., Templeton, W. J. F., Dugast, F. X., Narra, S. P. & To, A. C. Accelerating high-fidelity thermal process simulation of laser powder bed fusion via the computational fluid dynamics imposed finite element method (CIFEM). *Addit. Manuf. Lett.* 3, 100081 (2022).
- 29. Hemmasian, A. et al. Surrogate modeling of melt pool temperature field using deep learning. *Addit. Manuf. Lett.* **5**, 100123 (2023).
- Rodgers, T. M., Madison, J. D. & Tikare, V. Simulation of metal additive manufacturing microstructures using kinetic Monte Carlo. *Comput. Mater. Sci.* 135, 78–89 (2017).
- Yu, Y., Li, Y., Lin, F. & Yan, W. A multi-grid cellular automaton model for simulating dendrite growth and its application in additive manufacturing. *Addit. Manuf.* 47, 102284 (2021).
- Lian, Y., Lin, S., Yan, W., Liu, W. K. & Wagner, G. J. A parallelized threedimensional cellular automaton model for grain growth during additive manufacturing. *Comput. Mech.* **61**, 543–558 (2018).
- Yang, M., Wang, L. & Yan, W. Phase-field modeling of grain evolutions in additive manufacturing from nucleation, growth, to coarsening. *npj Comput. Mater.* 7, 56 (2021).
- Popova, E. et al. Process-structure linkages using a data science approach: application to simulated additive manufacturing data. *Integr. Mater. Manuf. Innov.* 6, 54–68 (2017).
- Ataollahi, S., Mahtabi, M., Yadollahi, A. & Mahtabi, M. J. Computational modeling of the effects of process parameters on the grain morphology of additively manufactured stainless steel. *Int. J. Adv. Manuf. Technol.* **125**, 3513–3526 (2023).
- Xue, T., Gan, Z., Liao, S. & Cao, J. Physics-embedded graph network for accelerating phase-field simulation of microstructure evolution in additive manufacturing. *npj Comput. Mater.* 8, 201 (2022).

- Choi, J. Y., Xue, T., Liao, S. & Cao, J. Accelerating phase-field simulation of three-dimensional microstructure evolution in laser powder bed fusion with composable machine learning predictions. *Addit. Manuf.* **79**, 103938 (2024).
- Sun, Z., Tsai, S.-P., Konijnenberg, P., Wang, J.-Y. & Zaefferer, S. A large-volume 3D EBSD study on additively manufactured 316L stainless steel. *Scr. Mater.* 238, 115723 (2024).
- Zhang, J., Wang, P. & Gao, R. X. Deep learning-based tensile strength prediction in fused deposition modeling. *Comput. Ind.* 107, 11–21 (2019).
- Marmarelis, M. G. & Ghanem, R. G. Data-driven stochastic optimization on manifolds for additive manufacturing. *Comput. Mater. Sci.* 181, 109750 (2020).
- Ponticelli, G. S., Giannini, O., Guarino, S. & Horn, M. An optimal fuzzy decision-making approach for laser powder bed fusion of AlSi10Mg alloy. *J. Manuf. Process.* 58, 712–723 (2020).
- Deb, K., Pratap, A., Agarwal, S. & Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evolut. Comput.* 6, 182–197 (2002).
- Hertlein, N., Deshpande, S., Venugopal, V., Kumar, M. & Anand, S. Prediction of selective laser melting part quality using hybrid Bayesian network. *Addit. Manuf.* 32, 101089 (2020).
- Xie, X. et al. Mechanistic data-driven prediction of as-built mechanical properties in metal additive manufacturing. *npj Comput. Mater.* 7, 86 (2021).
- 45. Fang, L. et al. Data-driven analysis of process, structure, and properties of additively manufactured Inconel 718 thin walls. *npj Comput. Mater.* **8**, 126 (2022).
- Zhang, M. et al. High cycle fatigue life prediction of laser additive manufactured stainless steel: a machine learning approach. *Int. J. Fatigue* **128**, 105194 (2019).
- Zhan, Z. & Li, H. Machine learning based fatigue life prediction with effects of additive manufacturing process parameters for printed SS 316L. *Int. J. Fatigue* **142**, 105941 (2021).
- Kusano, M. et al. Tensile properties prediction by multiple linear regression analysis for selective laser melted and post heat-treated Ti-6Al-4V with microstructural quantification. *Mater. Sci. Eng.: A* 787, 139549 (2020).
- Herriott, C. et al. A multi-scale, multi-physics modeling framework to predict spatial variation of properties in additive-manufactured metals. *Model. Simul. Mater. Sci. Eng.* 27, 025009 (2019).
- Herriott, C. & Spear, A. D. Predicting microstructure-dependent mechanical properties in additively manufactured metals with machineand deep-learning methods. *Comput. Mater. Sci.* **175**, 109599 (2020).
- Hong, R. et al. Artificial neural network-based geometry compensation to improve the printing accuracy of selective laser melting fabricated sub-millimetre overhang trusses. *Addit. Manuf.* 37, 101594 (2021).
- Hu, D., Grilli, N., Wang, L., Yang, M. & Yan, W. Microscale residual stresses in additively manufactured stainless steel: computational simulation. *J. Mech. Phys. Solids* **161**, 104822 (2022).
- 53. Chen, W. et al. Microscale residual stresses in additively manufactured stainless steel. *Nat. Commun.* **10**, 4338 (2019).
- Chia, H. Y., Wang, L. & Yan, W. Influence of oxygen content on melt pool dynamics in metal additive manufacturing: high-fidelity modeling with experimental validation. *Acta Mater.* 249, 118824 (2023).
- Giam, A., Chen, F., Cai, J. & Yan, W. Factorial design analytics on effects of material parameter uncertainties in multiphysics modeling of additive manufacturing. *npj Comput. Mater.* 9, 51 (2023).
- 56. Wang, Z. et al. Uncertainty quantification and reduction in metal additive manufacturing. *npj Comput. Mater.* **6**, 175 (2020).
- Ren, Z., Zhang, Z., Tian, Y. & Li, J. CRESt–Copilot for real-world experimental scientist. Preprint at https://doi.org/10.26434/ chemrxiv-2023-tnz1x (2023).

- Makatura, L. et al. How can large language models help humans in design and manufacturing? Part 1: Elements of The LLM-Enabled Computational Design and Manufacturing Pipeline. Harvard Data Science Review, (Special Issue 5). https://doi.org/10.1162/99608f92. cc80fe30 (2024).
- Touvron, H. et al. LLaMA: open and efficient foundation language models. Preprint at https://arxiv.org/abs/2302.13971 (2023).
- 60. Simonds, B. J. et al. Ability to simulate absorption and melt pool dynamics for laser melting of bare aluminum plate: Results and insights from the 2022 Asynchronous AM-Bench challenge. *Integr. Mater. Manuf. Innov.* **13**, 175–184 (2024).
- Kuan, A. et al. A common data dictionary and common data model for additive manufacturing. *Integr. Mater. Manuf. Innov.* 13, 105–119 (2024).
- 62. Wilkinson, M. D. et al. The FAIR guiding principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).
- 63. Roh, B.-M., Kumara, S. R., Witherell, P. & Simpson, T. W. Ontologybased process map for metal additive manufacturing. *J. Mater. Eng. Perform.* **30**, 8784–8797 (2021).
- Ko, H., Witherell, P., Lu, Y., Kim, S. & Rosen, D. W. Machine learning and knowledge graph based design rule construction for additive manufacturing. *Addit. Manuf.* 37, 101620 (2021).
- Sanfilippo, E. M., Belkadi, F. & Bernard, A. Ontology-based knowledge representation for additive manufacturing. *Comput. Ind.* 109, 182–194 (2019).
- Li, S., Feng, S., Kuan, A. & Lu, Y. Toward a standard data architecture for additive manufacturing. *JOM* 76, 1905–1912 (2024).
- Zhu, J. et al. A review of topology optimization for additive manufacturing: status and challenges. *Chin. J. Aeronaut.* 34, 91–110 (2021).
- Merchant, A. et al. Scaling deep learning for materials discovery. *Nature* 624, 80–85 (2023).
- Gradl, P. R. & Protz, C. S. Technology advancements for channel wall nozzle manufacturing in liquid rocket engines. *Acta Astronaut.* **174**, 148–158 (2020).
- 70. Ren, Z. et al. Machine learning-aided real-time detection of keyhole pore generation in laser powder bed fusion. *Science* **379**, 89–94 (2023).
- Wang, S., Zhu, L., Fuh, J. Y. H., Zhang, H. & Yan, W. Multi-physics modeling and Gaussian process regression analysis of cladding track geometry for direct energy deposition. *Opt. Lasers Eng.* **127**, 105950 (2020).

Acknowledgements

The authors acknowledge the financial support by the Ministry of Education, Singapore, under its Academic Research Fund Tier 2 (MOE-T2EP50221-0013). Z.H. acknowledges the support by the National University of Singapore through the President's Graduate Fellowship.

Author contributions

Z.H. wrote a major part of the manuscript. W.Y. conceived the idea. Both authors (Z.H. and W.Y.) reviewed and edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Wentao Yan.

Reprints and permissions information is available at http://www.nature.com/reprints

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/bync-nd/4.0/.

© The Author(s) 2024