

LayupFormer: A deep generative model for composite laminate layup design

Wenjie Xie, Hasan Caglar , Kailun Deng, David Ayre , Yifan Zhao ^{*}

Faculty of Engineering and Applied Sciences, Cranfield University, Cranfield, MK43 0AL, UK

ARTICLE INFO

Keywords:

Composite laminate design
Transformer
Tsai's invariants
Inverse design
Interpretability

ABSTRACT

Conventional laminate layup design relies on search-and-evaluate strategies that become intractable as ply counts grow, offering limited guarantees of feasibility, interpretability, and efficiency. While surrogate modelling and optimisation using artificial intelligence have accelerated composite design, most approaches still explore the design space via a search-centric manner. Generative methods offer an alternative by incorporating performance criteria, but often operate with restricted orientation sets, lack interpretability, or require auxiliary tools to ensure feasibility. This article introduces LayupFormer, a physics-informed Transformer framework that reformulates laminate layup design as an inverse sequence generation problem. The solution embeds mechanics through laminate parameters derived from Tsai's invariants and constrains designs with a domain-specific grammar over ply orientations. The framework couples a high-fidelity predictor, which regresses load and stiffness, with a generator that directly produces requirement-compliant layups efficiently across data scales. Attention analyses reveal that LayupFormer captures long-range through-thickness interactions and internalises laminate principle, providing interpretable insights into the generation process. Experimental validation confirms that LayupFormer-designed layups achieve superior bearing performance and reduced variability compared with empirical baselines. LayupFormer establishes a unified physics-informed generative framework that transforms laminate design from search-based optimisation into an interpretable and data-efficient inverse-design process, paving the way for scalable and automated composite design.

Nomenclature

Roman symbols	
A	Extensional stiffness matrix of the laminate
B	Bending-extension coupling stiffness matrix of the laminate
D	Bending stiffness matrix of the laminate
t	The specific t-th layer of the laminate
T	Total number of plies in the laminate
P	Maximum bearing load in the bearing test
K	Initial bearing stiffness obtained from the load-displacement response
Abbreviations	
ASTM	American Society for Testing and Materials
ASTM D5961	Standard Test Method for Bearing Response of Polymer Matrix Composite Laminates
M	Composite Laminates
VA-RTM	Vacuum-assisted resin transfer moulding
FEM	Finite element method
Abaqus	Commercial finite element software used for high-fidelity simulations
DD	Double-double laminate family

(continued on next column)

(continued)

Quad	Quasi-isotropic laminate family
Sym	Symmetric laminate family
CLT	Classical laminate theory
DSL	Domain-specific language
CNN	Convolutional neural network
ANN	Artificial neural network
GAN	Generative adversarial network
ML	Machine learning
LayupFormer	Deep generative model proposed in this work for laminate layup design

1. Introduction

Fibre-reinforced polymer (FRP) composites are increasingly employed in aerospace, wind energy, marine, and civil engineering due to their high specific strength, fatigue resistance, and design flexibility

^{*} Corresponding author.

E-mail address: yifan.zhao@cranfield.ac.uk (Y. Zhao).

[1,2]. Among them, laminated composites are particularly attractive because their mechanical properties can be tailored through ply orientations and stacking sequences, enabling near-optimal performance under diverse service conditions [3,4]. However, this designability also renders the layup design problem highly complex, as small variations in ply angles or stacking order can result in significant changes in stiffness, strength, and failure behaviour. The challenge intensifies as the number of plies grows, leading to an exponentially expanding solution space [5].

Classical laminate theory (CLT) provides an analytical framework for characterising laminate stiffness [6,7]. Building on CLT, Tsai and co-workers proposed invariant-based descriptions of laminate stiffness and the Double-Double (DD) family, offering new analytical design routes [8–11]. While these approaches provide fast analytical insights, they cannot capture realistic load bearing and failure processes under complex service conditions, necessitating more sophisticated optimisation strategies. Meta-heuristic optimisation methods such as genetic algorithms (GA), ant colony optimisation (ACO) [12–16], and branch-and-bound have been combined with CLT to search for optimal stacking sequences [17,18]. To further enhance efficiency, 2D and 3D sampling optimisation methods have been developed, which utilise continuous lamination parameter feasible regions to robustly guide the design [19–22]. However, even these advanced strategies typically rely on a multi-step retrieval process from continuous parameters to discrete layers, essentially remaining bound to the iterative search framework. More recently, machine learning (ML) surrogates trained on finite element analysis (FEA) data have accelerated laminate design by providing rapid performance evaluation [4,23–25]. Nevertheless, these methods still follow a forward design paradigm: propose a laminate, evaluate its performance, and iteratively refine via optimisation. Such forward approaches face two major challenges: (i) the search process is only loosely aligned with laminate design logic, and (ii) the solution space expands exponentially with ply count, making optimisation computationally intractable for realistic laminates [26–29]. Generative methods offer an alternative inverse design route. Conditional GANs (cGANs), for instance, have been applied to generate laminates that satisfy given requirements [30–32]. However, their application is typically limited to restricted laminate families (e.g., Quad or DD types) and their stochastic generation lacks interpretability, obscuring the relationship between requirements and layups.

Inspired by recent advances in natural language processing (NLP), where Transformer-based generative models have achieved remarkable success in sequence generation tasks [33–36], laminate layup design can be reformulated as a requirement-conditioned sequence generation problem. In this view, ply orientations and laminate types constitute a domain-specific language (DSL) governed by syntactic rules, while performance requirements are embedded into high-dimensional latent representations. This analogy enables the development of generative frameworks that directly map design requirements to feasible and manufacturable layup sequences, bridging the gap between optimisation-based search and interpretable inverse design.

In this paper, we propose LayupFormer, a physics-informed Transformer framework for composite laminate inverse design. LayupFormer integrates a predictor with Tsai's modulus-informed embedding to ensure physics-consistent performance regression, and a grammar-constrained generator for requirement-conditioned layup sequence generation. LayupFormer is validated on an open-hole tensile case but is not limited to this scenario; its design principles are extensible to diverse objectives, laminate families, and loading conditions. The contributions of this work are as follows:

- Reformulation of the laminate layup design as a requirement-conditioned sequence generation task inspired by NLP.
- Introduction of a physics-informed embedding derived from Tsai's invariants to enhance sample efficiency, training stability, and interpretability.

- Development of a Transformer-based predictor-generator framework (LayupFormer) that couples physics consistency with attention-based interpretability.
- Establishment of a closed-loop evaluation pipeline integrating FEM data generation, large-scale surrogate expansion, and experimental tests.

The organisation of this article is presented below. Section 2 introduces the proposed methodology, including the formulation of a domain-specific language for laminate design, the architecture of the LayupFormer predictor and generator, and the evaluation and validation protocols. Section 3 reports the results obtained in the open-hole tensile case, followed by an in-depth discussion of physics-informed embedding, attention-based interpretability, and validation scope in Section 4. Section 5 concludes the paper with a summary of the key research findings and potential directions for future work.

2. Methodology

The methodology is presented from three perspectives:

- The design of a domain-specific language (DSL) that encodes laminate layup sequences and their governing grammar rules.
- The architecture of the LayupFormer framework, which consists of a physics-informed predictor and a grammar-constrained generator.
- The evaluation and validation protocols, which define the case study and the procedures used to assess the effectiveness of the framework.

2.1. Data pattern: domain-specific language for layup design

The hypothesis of this study is that laminate layup design can be reformulated as a rule-constrained sequence generation problem, where ply orientations form a structured sequence subject to manufacturing and mechanical rules. This sequential nature closely resembles how words in natural language combine under syntactic and semantic constraints to form meaningful sentences.

To illustrate this analogy, Fig. 1 presents the structural correspondence between laminate layup design and language modelling. Each ply orientation is represented as a token, while stacking rules such as balance and symmetry act as grammatical constraints that define feasible combinations. Performance requirements play a role analogous to linguistic context, conditioning the generation process toward valid and meaningful sequences. Both problems therefore share the nature of rule-constrained sequence generation.

Building on this analogy, we define a domain-specific language (DSL) that formally encodes laminate feasibility and manufacturability through explicit grammatical constraints. It comprises three key elements:

- Ply angle tokens: discrete orientations drawn from the range $[-85^\circ, 90^\circ]$ with a step of 5° , yielding 36 admissible values based on manufacturing precision.
- Laminate type token: indicating the structure class, namely Quad, DD, and Sym.
- Special tokens: including $\langle \text{bos} \rangle$, $\langle \text{eos} \rangle$, $\langle \text{pad} \rangle$, and $\langle \text{mask} \rangle$, required for Transformer-based autoregressive generation.

In addition to the lexicon, laminate design requires adherence to syntactic grammar rules:

- DD laminate: the first four plies form two $\pm\theta$ pairs, and the remaining plies are obtained by repeating this four-ply unit.

Both are rule-constrained sequence generation problems

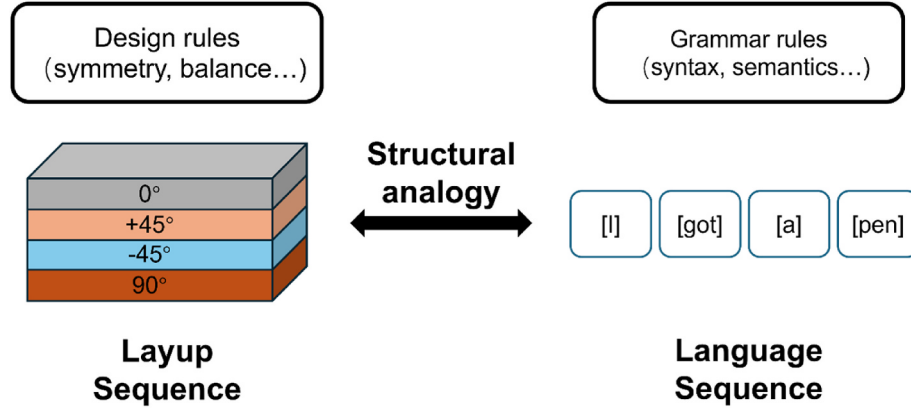


Fig. 1. Conceptual analogy between laminate layup design and language modelling.

- (ii) Quad laminate: the first half-stack is generated using only $[0^\circ, \pm 45^\circ, 90^\circ]$; the second half-stack mirrors the first. Balanced counts of $+45^\circ$ and -45° are enforced.
- (iii) Sym laminate: the first half-stack may include arbitrary orientations, but each non- $0^\circ/90^\circ$ angle must appear with its opposite to satisfy symmetry and balance; the second half-stack mirrors the first.

This DSL explicitly encodes laminate feasibility and manufacturability as grammatical constraints, ensuring that any generated sequence corresponds to a physically valid laminate. It provides the foundation upon which LayupFormer learns the mapping between design requirements and layup sequences and can be readily extended to accommodate different performance objectives, laminate families, and loading scenarios.

2.2. LayupFormer framework

The proposed LayupFormer is a physics-informed generative framework designed for inverse laminate layup design. It integrates two complementary modules: (i) a predictor, which establishes a physics-consistent mapping from ply sequences to laminate performance, and (ii) a generator, which realises the inverse process by producing requirement-compliant layups conditioned on target properties.

Together, these two modules form a closed generative-evaluative loop that enables direct mapping from performance requirements to feasible and manufacturable laminate designs. The overall architecture is illustrated in Fig. 2, where the generator employs a Transformer

encoder-decoder to autoregressively generate layup sequences under syntactic and physics-informed constraints, while the predictor employs an encoder-only Transformer to regress target properties from ply sequences.

The predictor incorporates Tsai's modulus-informed embedding, which grounds the input representation in laminate mechanics, thereby improving training stability, sample efficiency, and interpretability. Its trained model not only functions as a standalone regression tool but also serves as a surrogate evaluator in the generator's training. The generator, by contrast, conditions on design requirements (performance indicators and laminate type) and produces valid sequences guided by grammar rules and further constrained through performance-consistency checks against the predictor.

2.2.1. LayupFormer predictor

The predictor in LayupFormer establishes the forward mapping from laminate layup sequences to performance indicators, specifically the maximum load-bearing capacity P and the equivalent stiffness K . As illustrated in Fig. 3, it adopts an encoder-only Transformer architecture, designed to capture both local interactions and long-range dependencies across plies. Each encoder layer consists of multi-head self-attention, feed-forward sublayers, residual connections, and layer normalisation, enabling the model to internalise through-thickness interactions inherent in composite laminates.

To ground the predictor in physical principles, LayupFormer incorporates an embedding strategy based on Tsai's invariants [7,11]. For an orthotropic ply with reduced stiffness matrix Q , its transformed stiffness in the laminate coordinate system at orientation θ can be

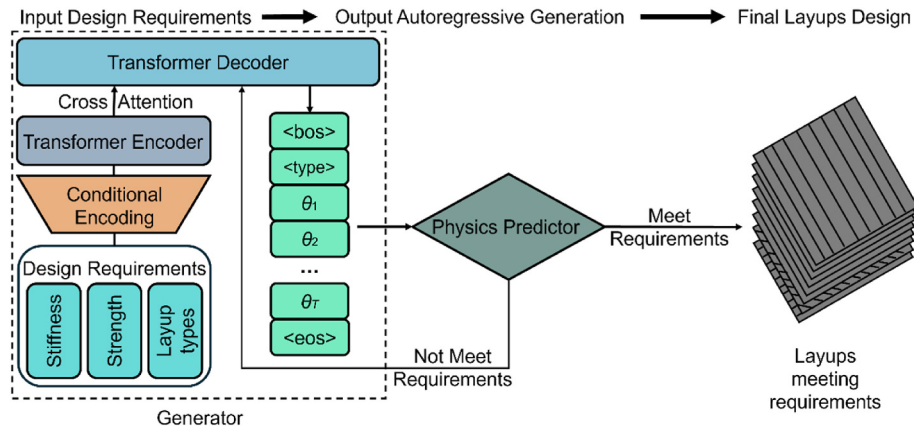


Fig. 2. The proposed LayupFormer architecture.

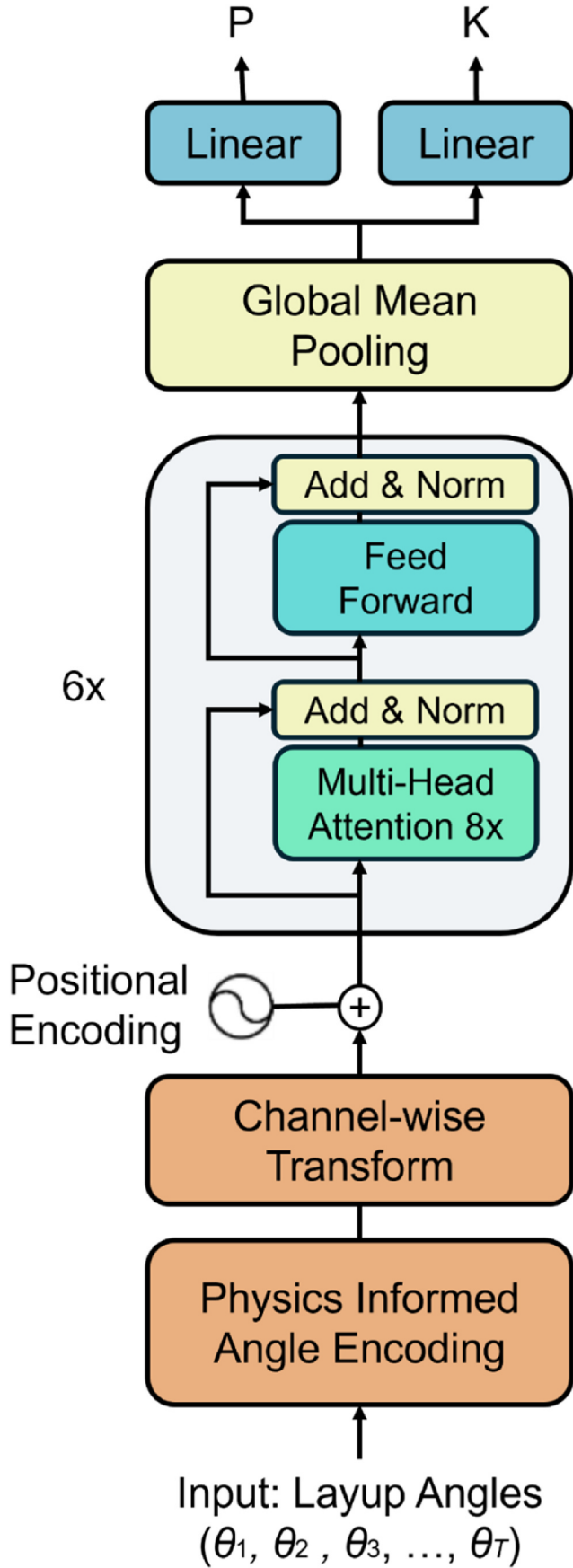


Fig. 3. Architecture of the LayupFormer predictor.

expressed as:

$$\bar{Q}(\theta) = \sum_{j=1}^4 I_j T_j(\theta), \quad (1)$$

where I_j are material invariants independent of θ , and $T_j(\theta)$ are trigonometric functions of the orientation angle. Integrating ply contributions through the thickness yields the laminate stiffness matrix ABD :

$$ABD = \sum_{k=1}^T (\bar{Q})_k \begin{bmatrix} (z_k - z_{k-1}) & \frac{1}{2} (z_k^2 - z_{k-1}^2) \\ \frac{1}{2} (z_k^2 - z_{k-1}^2) & \frac{1}{3} (z_k^3 - z_{k-1}^3) \end{bmatrix}, \quad (2)$$

where z denotes the coordinates in the thickness direction, T represents the total number of layers, and k refers to the k -th ply.

Drawing inspiration from Tsai's invariant formulation, we represent each ply angle θ by a harmonic feature vector directly linked to laminate parameters:

$$\Phi(\theta) = [\cos \theta, \sin \theta, \cos 2\theta, \sin 2\theta, \cos 4\theta, \sin 4\theta]. \quad (3)$$

Tsai showed that the components of the transformed ply stiffness and lamination parameters can be expressed as linear combinations of $\cos 2\theta$, $\sin 2\theta$, $\cos 4\theta$, $\sin 4\theta$, with coefficients given by material invariants. The feature vector $\Phi(\theta)$ therefore explicitly includes these fundamental stiffness harmonics ($2\theta, 4\theta$) and augments them with first-order terms ($\cos \theta, \sin \theta$) to preserve the geometric uniqueness of the ply orientation. For a laminate with T plies, the layout $\{\theta_k\}_{k=1}^T$ is encoded as a sequence of such harmonic embeddings, forming an input matrix $E \in \mathbb{R}^{T \times 6}$ with rows $E_k = \Phi(\theta_k)$. This sequence E serves as the tokenised input to the predictor, ensuring that each token carries an angle representation aligned with the spectral basis of Tsai's invariant theory rather than a raw scalar. In this way, the network does not need to rediscover the periodicity and symmetry of the ply orientations from data alone, which empirically improves the regression accuracy and ensures better generalisation to unseen layouts compared with angle-only or simple $[\cos \theta, \sin \theta]$ encodings.

Given a sequence of embedded ply orientations $\{\Phi(\theta_1), \Phi(\theta_2), \dots, \Phi(\theta_T)\}$, convolutional layers first extract local interactions, followed by a linear projection and positional encoding. This representation is then processed by a stack of Transformer encoder layers. A pooling operation aggregates the contextualised sequence embedding, which is passed through a shared projection layer and two regression heads to simultaneously predict P and K . It can be written as:

$$\hat{Y} = [\hat{P}, \hat{K}] = f_{\text{Transformer}}(\{\Phi(\theta_k)\}_{k=1}^T). \quad (4)$$

The predictor is trained in a multi-task regression setting with mean squared error loss applied to both targets, ensuring simultaneous learning of stiffness and strength indicators. In addition to its standalone role as a regression model, the trained predictor also functions as a frozen evaluator during the training of the generator, enforcing physics consistency in requirement-conditioned sequence generation.

While in this study the predictor is trained on 24-ply laminates and regresses the two case-specific indicators P and K , the architecture itself is not restricted to these choices. Any vectorizable design objective (e.g., buckling loads, fatigue resistance, multi-objective performance metrics) can serve as a target output without altering the network architecture. Similarly, the laminate thickness (ply count) is determined by the grammar rules defined in Section 2.1; by adjusting these rules, the same framework can handle variable ply numbers. Thus, the predictor should be viewed as a general physics-informed regression module.

2.2.2. LayupFormer generator

While the predictor establishes the forward mapping from layup sequences to performance, the generator realises the inverse process, producing layup sequences conditioned on design requirements. It

adopts a Transformer encoder–decoder architecture (as shown in Fig. 4), where the encoder embeds conditioning signals and the decoder autoregressively generates ply sequences under causal masking.

The generator conditions on continuous performance requirements (e.g., P, K) together with the discrete laminate type (Quad, DD, Sym). These inputs are embedded into latent vectors and provided as memory for cross-attention in the decoder. During autoregressive decoding, each new ply token is generated sequentially with strict causal masking, ensuring that the model only attends to past tokens and the conditioning memory, thereby maintaining consistency between training and inference.

To ensure manufacturability and feasibility, the generator incorporates the domain-specific grammar defined in Section 2.1 as constraints during decoding. At each step, infeasible tokens are masked out based on laminate type, and lightweight syntactic validation is performed to prevent rule violations. For example, in DD laminates only the

first four plies are generated and subsequently repeated, while in Quad and Sym laminates only the first half-stack is generated and mirrored to complete the full sequence. This integration of grammar rules ensures that all generated sequences are syntactically valid by construction.

The generator is trained under an autoregressive maximum-likelihood framework; however, likelihood alone is insufficient to guarantee requirement-compliant laminates. To address this limitation, two additional supervisory signals are incorporated. The first is a performance consistency term, which encourages generated sequences to remain aligned with the specified performance targets. The second is a physics-consistency term, which evaluates generated layups using the frozen predictor introduced in Section 2.2.1 and penalises deviations from the required mechanical behaviour. Together, these auxiliary signals constrain the generator to remain within the feasible mechanical domain, thereby enhancing the reliability of conditional layup generation compared with a purely likelihood-driven approach.

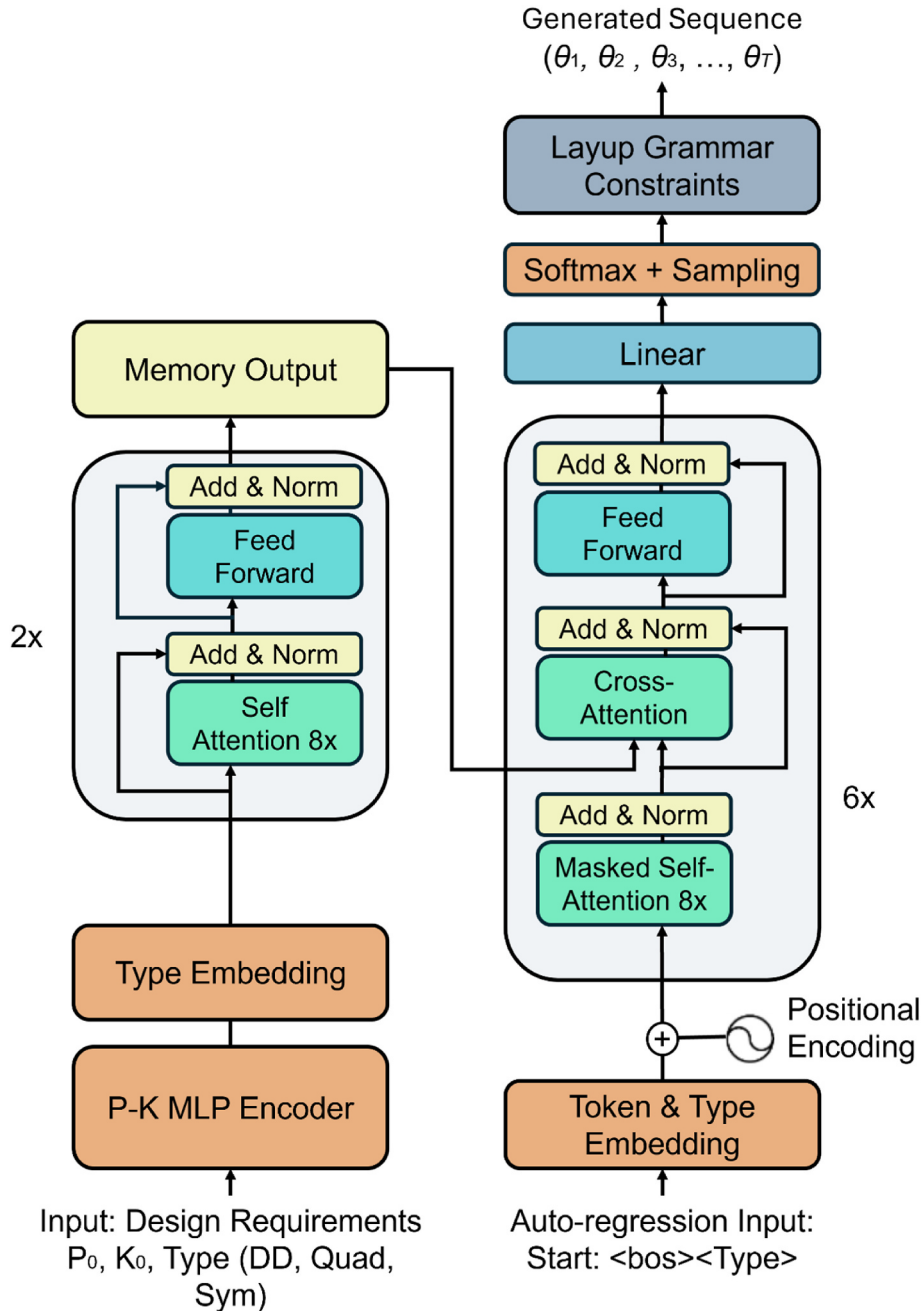


Fig. 4. Architecture of the LayupFormer generator.

In addition to improving feasibility, the attention mechanisms within the generator provide interpretability. Cross-attention explicitly reveals how conditioning requirements (performance indicators and laminate type) influence each decoding step, while masked self-attention traces how previously generated plies shape subsequent decisions. This offers mechanistic insights into the generation process, distinguishing LayupFormer from conventional black-box generative models.

Same as mentioned in Section 2.2.1, any design requirement that can be vectorised may serve as conditioning input, and variable ply counts can be accommodated by extending the grammar rules in Section 2.1. Thus, the generator should be viewed as a general conditional sequence model for laminate design.

2.3. Evaluation and validation

To assess the effectiveness of LayupFormer, a set of evaluation and validation protocols was established. These protocols serve two purposes: first, to verify the predictor's ability to regress laminate performance with high fidelity; and second, to evaluate the generator's capability to produce requirement-compliant, syntactically valid, and diverse laminates under realistic loading conditions. In addition, experimental validation was conducted to demonstrate the practical feasibility of the generated laminates.

The following subsections detail the case definition, data acquisition strategy, evaluation procedures, and experimental setup.

2.3.1. Problem definition and data acquisition

The evaluation was carried out on an open-hole laminate under tensile loading, a representative scenario for various engineering applications such as lug joints in aircraft structures, root regions of wind turbine blades, and bolted connections in rotor hubs. This configuration was selected because it couples local stress concentrations with global load transfer, making it a stringent benchmark for laminate design.

A total of 24 plies are considered in this study, consistent with industrial practice for balanced and quasi-symmetric laminates. While the framework itself is not restricted to a fixed ply count, this choice provides a realistic and challenging test case. The design requirements are formulated in terms of two performance indicators: the maximum load-bearing capacity (P) and the equivalent in-plane stiffness (K). These two metrics were chosen as representative structural indicators, but the framework is general and can incorporate any set of objectives that can be expressed as vectors.

High-fidelity training data were generated using finite element simulations implemented in Abaqus. The simulations employed a Hashin-based progressive damage model [37] with stiffness degradation to capture the nonlinear failure process. The high-fidelity simulations were carried out in Abaqus/Explicit 2022 using a shell representation of the laminate. Each ply was modelled with eight-node continuum shell elements with reduced integration (SC8R), stacked through the thickness according to the specified layup. A nominal in-plane mesh size of 3 mm was used in the far field, while the region around the hole was radially refined such that approximately 40 elements were placed along the hole circumference to capture the local bearing response. The steel pin was modelled as a rigid body. The boundary conditions were replicated from the pin-loaded bearing test. On the fixed side, the nodes on the laminate surface in contact with the pin were kinematically coupled to an external reference point, which was fully constrained to prevent rigid-body motion. On the loaded side, the bearing load was applied as a prescribed displacement at the centre of the rigid pin. This setup avoids artificial point loading on individual nodes and ensures a realistic transfer of load between the rigid pin and the composite laminate. The FEM model reproduces the specimen geometry and boundary conditions of the ASTM D5961M [38] bearing test for polymer-matrix composite laminates.

For each layup sampled under the grammar constraints defined in Section 2.1, the maximum load P was extracted from the failure point of

the load-displacement response, while the equivalent stiffness K was computed from the slope of the initial elastic regime. In total, 4000 distinct layup configurations were analysed using high-fidelity FEM in Abaqus to provide ground-truth bearing load and stiffness pairs for training the predictor. All simulations and network training in this study were carried out on a laptop equipped with an Intel Core i7-9750H CPU, 32 GB RAM and an NVIDIA RTX 2070 GPU with 8 GB memory. Generating the 4000 high-fidelity samples and training the predictor required approximately 26 days in total on this machine. Once a high-accuracy predictor had been obtained, it was used to augment the dataset via fast surrogate evaluations for training the generator. The problem setting and data generation workflow are illustrated in Fig. 5. Panel (a) shows the geometry of the open-hole laminate, (b) presents the boundary conditions and loading configuration, and (c) summarises the automated data acquisition pipeline combining laminate sampling, FEM simulation, and performance extraction.

2.3.2. Predictor evaluation

The predictor was trained and evaluated on the FEM-generated dataset described in Section 2.3.1. To ensure statistical reliability, the dataset was randomly divided into 80 % training, 10 % validation, and 10 % test splits. The validation set was used for early stopping and learning-rate scheduling, while final performance was reported on the independent test set.

Prior to training, all performance indicators were standardised. Given a dataset of layup-performance pairs $\{(x_i, y_i)\}_{i=1}^N$, where x_i denotes the sequence of ply orientations and $y_i = [P_i, K_i]$ represents the corresponding target properties, the standardisation yields

$$\tilde{y}_i = \frac{y_i - \mu}{\sigma}, \quad (5)$$

with μ and σ denoting the mean and standard deviation of each indicator over the training set.

The predictor outputs a pair of predicted values $\hat{y}_i = [\hat{P}_i, \hat{K}_i]$. The training objective is formulated as the mean squared error (MSE) is applied to both targets:

$$\mathcal{L}_{pred} = \frac{1}{N} \sum_{i=1}^N ((\hat{P}_i - \tilde{P}_i)^2 + (\hat{K}_i - \tilde{K}_i)^2). \quad (6)$$

This formulation ensures that both stiffness and strength indicators are learned simultaneously, without requiring manual weighting, since both targets are normalised to the same scale. Training was performed using the AdamW optimiser with learning-rate scheduling, and batch normalisation was applied within the convolutional layers to stabilise convergence. Early stopping was employed based on the validation loss to prevent overfitting.

To evaluate predictive accuracy, the predictor was assessed on the independent test set using the coefficient of determination (R^2) for each indicator, reported in the original (de-standardised) space. In addition, attention maps from the final encoder layer were analysed to examine how the model aggregates through-thickness information, providing interpretability alongside numerical accuracy.

To systematically examine the effect of input representation and network architecture, the predictor was compared against baseline models under different settings. Three encoding strategies were tested: (i) raw ply angles, (ii) sine-cosine embedding, and (iii) the proposed Tsai-invariant embedding. Each encoding was evaluated with three model families: artificial neural networks (ANNs), convolutional neural networks (CNNs), and the proposed LayupFormer predictor. This design enables disentangling the influence of input representation from that of model architecture and provides a structured basis for the comparative results presented in Section 3.

2.3.3. Generator evaluation

The generator was trained on the large-scale dataset produced by the

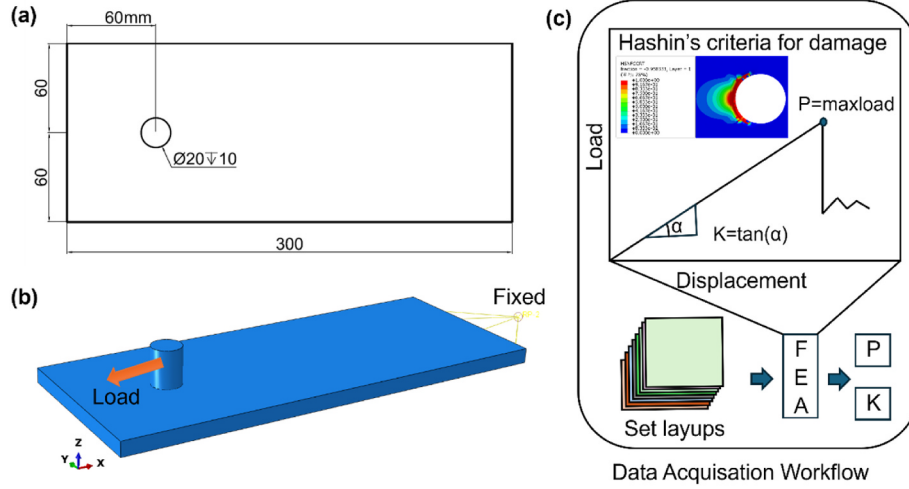


Fig. 5. The process of generating training data using FEM. (a) part dimensions. (b) boundary conditions and loading configuration. (c) data acquisition workflow with Hashin's criterion.

well-trained predictor, which provided paired layup sequences and corresponding performance indicators. Since Transformer-based sequence decoders typically require sufficient data to achieve stable convergence, three dataset scales were considered: 24k, 240k, and 1.4M samples. Early stopping with a patience of ten epochs was applied to avoid overfitting.

The training objective of the generator extends the standard autoregressive maximum-likelihood framework by introducing auxiliary supervision to strengthen conditional generation. Specifically, the total loss function combines three complementary objectives:

$$\mathcal{L}_{gen} = \mathcal{L}_{CE} + \lambda_{perf} \mathcal{L}_{perf} + \lambda_{phys} \mathcal{L}_{phys}, \quad (7)$$

where $\mathcal{L}_{CE}$ denotes the standard cross-entropy loss under teacher forcing, under teacher forcing, ensuring syntactic correctness and stable sequence learning; $\mathcal{L}_{perf}$ represents a performance alignment loss of obtained from a lightweight regression head attached to the decoder's shared representation, which computes a mean-squared error between the predicted and target performance indicators, thereby guiding the hidden representations toward numerical consistency with the specified requirements; and $\mathcal{L}_{phys}$ is the physics-consistency loss, computed by decoding candidate sequences, evaluating its predicted performance using the frozen predictor, and applying a *ReLU* penalty so that only underperforming cases are penalised. In this way, λ_{perf} encourages quantitative alignment with the conditioning targets, while $\mathcal{L}_{phys}$ enforces feasibility within the mechanical domain. The weighting factors λ_{perf} and λ_{phys} were empirically chosen to balance the learning of grammatical validity, performance alignment, and physics consistency, here $\lambda_{perf} = 0.2$, $\lambda_{phys} = 0.3$.

Evaluation of the generator focused on three aspects. Syntactic validity was measured by the proportion of generated layups satisfying the grammar rules defined in Section 2.1. Performance alignment was quantified by the relative error between the generated layup properties, evaluated via the frozen predictor, and the target requirements. Finally, design diversity was assessed by the uniqueness ratio of generated sequences and the distributional spread of ply orientations. Together, these metrics provide a comprehensive assessment of whether the generator can produce laminates that are valid, requirement-compliant, and sufficiently diverse to support exploratory design.

2.3.4. Experiment validation

To validate the practical applicability of LayupFormer, representative laminates were manufactured and tested under the same loading condition as defined in Section 2.3.1. The FEM dataset was generated

using the T700GC/M21 system, owing to the availability of detailed mechanical and chemical parameters, which are required for reliable modelling. For the physical experiments, however, the Zoltek PX35/epoxy system was adopted. This material was chosen because it offers stable processing characteristics under laboratory-scale manufacturing while exhibiting mechanical performance broadly consistent with T700GC/M21.

The use of two different material systems is justified on theoretical grounds. According to Tsai's invariant formulation, laminate performance can be expressed as a product of two independent components: material invariants and laminate parameters. This decoupling implies that, for a fixed geometry and loading condition, the influence of ply stacking on performance is independent of the specific material system. Consequently, the design principles and layup configurations proposed by LayupFormer are transferable across materials. In this work, Zoltek PX35/epoxy was therefore employed for manufacturing convenience, while T700GC/M21 was retained in modelling due to the completeness of its published property data.

Panels were fabricated via vacuum-assisted resin transfer moulding (VA-RTM) and subsequently cut into coupons following the ASTM D5961M [38]. Each coupon had an overall length of 300 mm, a width of 120 mm, and a nominal thickness of 10 mm. A 20 mm diameter open hole was drilled on the specimen centreline at an edge distance of 60 mm from the loaded end. The same geometry was used as in the FEM model. Two layup configurations were prepared: an empirical Quad baseline widely used in practice and a Quad sequence proposed by LayupFormer as a candidate with comparable predicted performance. Mechanical testing was performed using a standard pin-loaded fixture in accordance with ASTM D5961M. For each configuration, three specimens were tested, and the resulting load-displacement curves were recorded. From these responses, the nominal bearing stress and initial bearing stiffness were derived following the data reduction procedure in ASTM D5961M.

The experimental setup and specimens are illustrated in Fig. 6. Fig. 6 (a) shows the pin-loaded test fixture used for ASTM D5961M bearing tests. Fig. 6 (b) presents the manufactured specimens prior to testing, while Fig. 6 (c) displays representative specimens after failure, revealing typical damage modes under bearing loads.

This experimental campaign aims to validate the LayupFormer framework by examining the structural performance of its generated layups across different material systems, providing a basis for evaluating the method's generality and practical applicability.

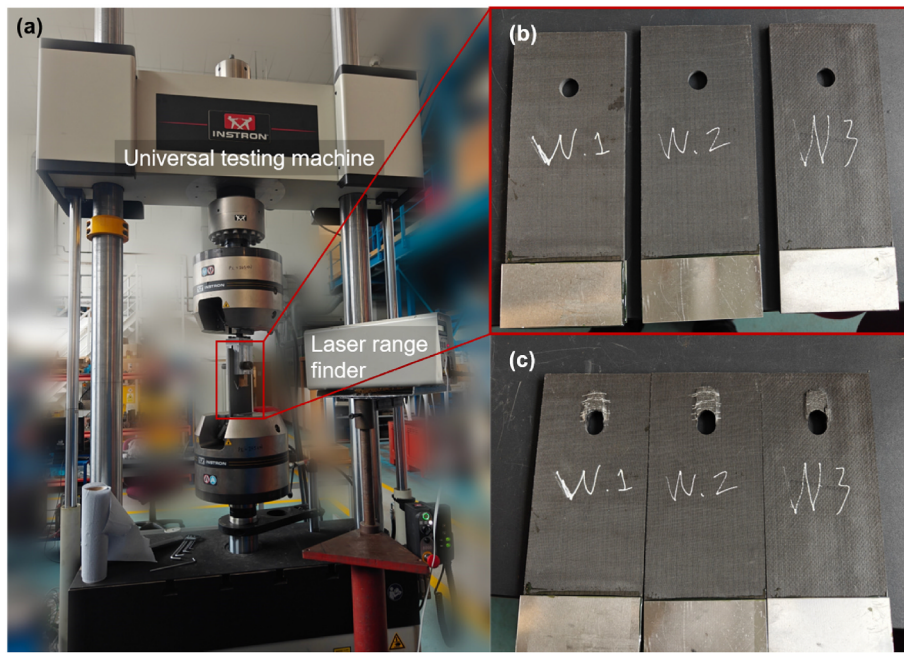


Fig. 6. (a) Test apparatus, (b) Specimens before testing, (c) Specimens after testing.

3. Results

This section presents the evaluation of LayupFormer on the open-hole tensile case introduced in Section 2.3.1. The results are reported in four parts. Section 3.1 examines the performance of the predictor, demonstrating its regression accuracy and its capability to generate large-scale surrogate data for subsequent use. Section 3.2 evaluates the generator, with particular focus on training behaviour, the effect of dataset scale, and the trade-off between validity, alignment, and diversity. Section 3.3 analyses the attention maps of both the predictor and the generator, exploring the interpretability offered by the framework. Finally, Section 3.4 reports the experimental validation, where LayupFormer-generated layups are compared against an empirical baseline under ASTM D5961M bearing tests. Together, these results establish the effectiveness, interpretability, and practical feasibility of the proposed framework.

3.1. Predictor evaluation

The predictive performance of LayupFormer was first assessed on the high-fidelity FEM dataset described in Section 2.3.1. Fig. 7 compares the ground truth and the predicted values of P and K under three different input representations. These include raw ply angles, sine-cosine embedding and the proposed physics-informed representation. The evaluation covers ANN, CNN and Transformer models. Across all tested settings, the Transformer predictor equipped with the physics-informed embedding delivers the highest accuracy, with predicted values closely matching the reference data. This confirms that embedding laminate parameters derived from Tsai's invariants improves model training, particularly when only limited data are available.

The trained predictor was then used as a surrogate to generate a large-scale dataset for generator training. Fig. 8 illustrates the structure of this dataset. Fig. 8 (a) shows a t-SNE projection of 50,000 randomly sampled layups where three well-separated clusters correspond to Quad, DD and Sym laminate types. The clear separation indicates that the surrogate-generated dataset preserves distinct structural characteristics among laminate families, confirming that the domain-specific language and grammar constraints effectively encode type-dependent design features. Fig. 8(b) and (c) illustrate the distribution ranges of the two

performance indicators, P and K, stratified by laminate type. The DD family exhibits the widest variability, the Quad family is the most concentrated, and the Sym family shows the highest median values. These patterns reflect distinct structural characteristics among laminate families while indicating that the surrogate-generated dataset spans a broad and balanced performance range. Such diversity in P-K coverage ensures sufficient variability and representativeness for robust generator training.

3.2. Generator evaluation

The performance of the LayupFormer generator was evaluated using the surrogate datasets produced by the trained predictor. To examine the influence of dataset scale on training behaviour, three dataset sizes were considered, namely 24k, 240k and 1.4M layup-performance pairs. Each dataset was sampled with consistent proportions of DD, Quad and Sym laminates, and with similar P-K ranges to ensure comparability across scales.

Table 1 and Fig. 9 summarise the training behaviour across the three dataset scales. The convergence curves in Fig. 9 demonstrate that the generator achieves stable optimisation in all cases. Although the total loss values are comparable, the breakdown in Table 1 reveals distinct trends among the three components. The CE loss remains relatively moderate and nearly invariant with dataset scale, which is expected because the mapping from performance requirements to layup sequences is inherently one to many, meaning that multiple valid laminates can satisfy the same target. As a result, the model learns a probabilistic distribution rather than converging to a single deterministic solution, leading to a non-negligible residual CE. The supervised loss exhibits a modest decrease with increasing dataset scale, indicating improved conditional alignment, whereas the physics-consistency loss remains largely stable, suggesting that physical feasibility is effectively maintained across scales. These results indicate that the training strategy is robust across data scales, with larger datasets enhancing conditional alignment while maintaining consistent physics-based feasibility.

The effect of dataset scale on generation performance and diversity is presented in Fig. 10. In terms of generation performance (Fig. 10(a)), it was revealed that larger datasets lead to higher success rates and lower average errors across all laminate types, indicating improved alignment

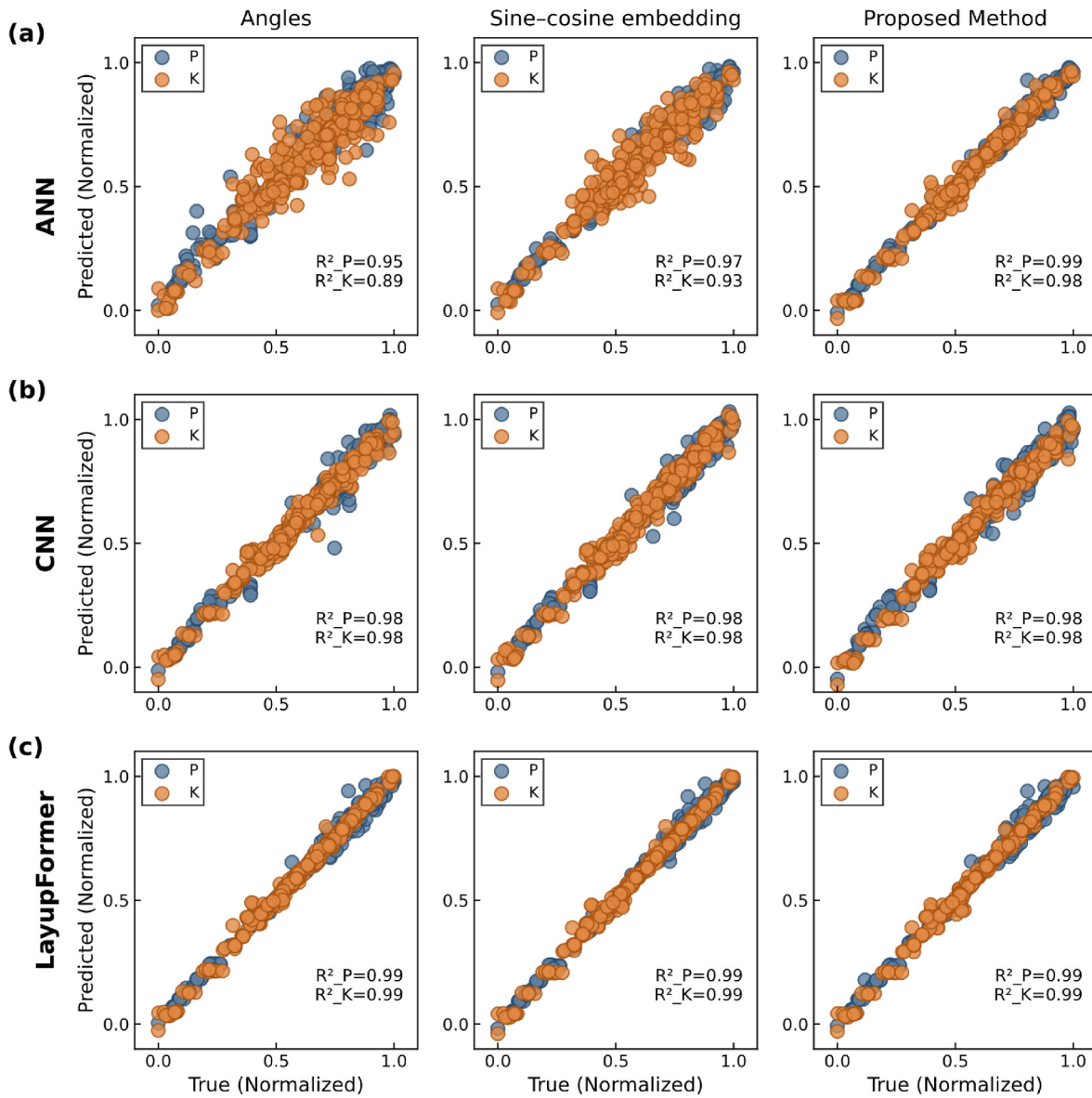


Fig. 7. Ground truth (x) vs. predicted value (y) under different representation schemes: (a). ANNs predictor, (b). CNNs predictor, (c). LayupFormer predictor.

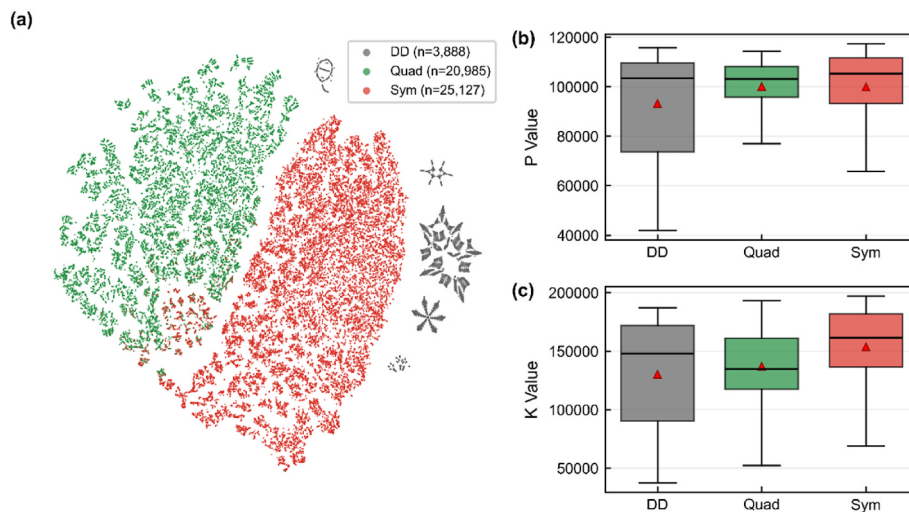


Fig. 8. (a) t-SNE visualisation of the sampled training data. (b) distribution of P in the training data. (c) distribution of K in the training data.

Table 1
Final loss values for different training data scale.

Data Scale	Total Loss	CE Loss	Supervised Loss	Physics Loss
24k	0.757474	1.197422	0.038284	0.401599
240k	0.781579	1.255114	0.022844	0.437259
1.4M	0.755018	1.252524	0.014694	0.380209

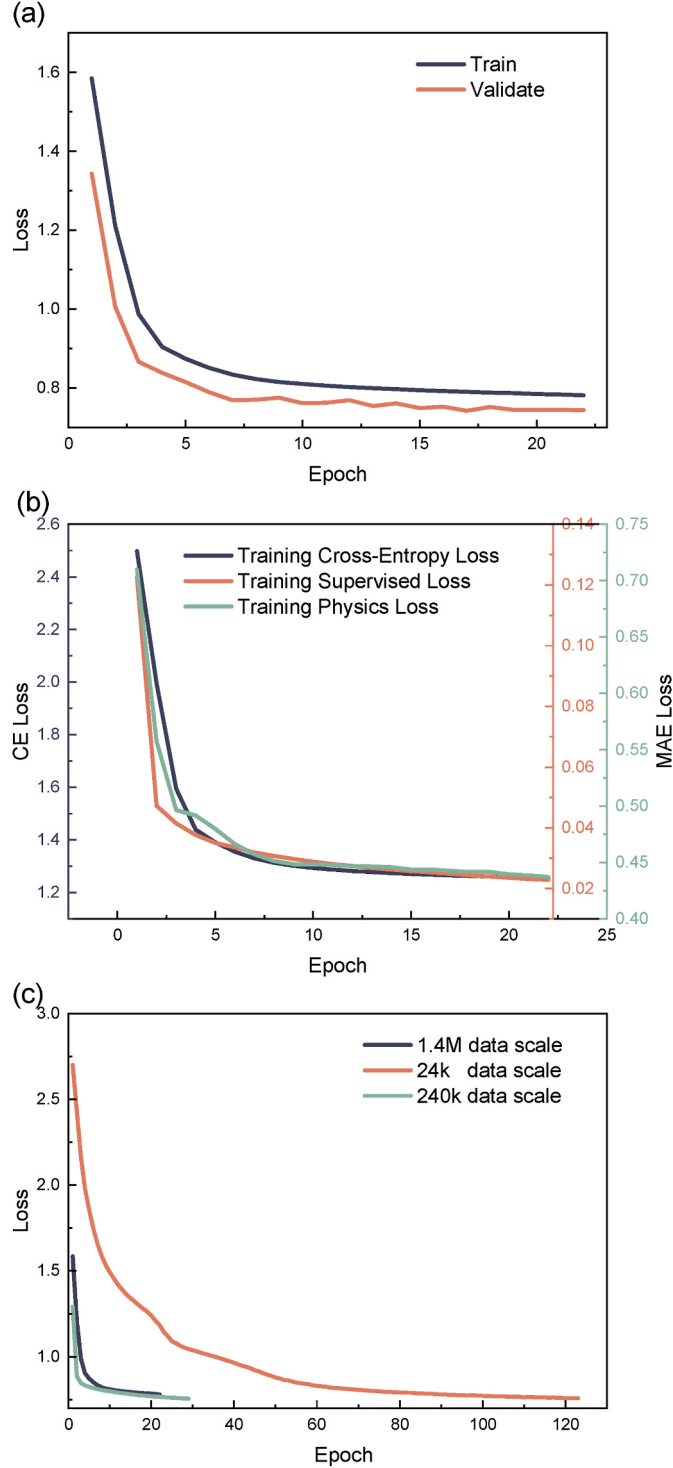


Fig. 9. Training behaviour of LayupFormer generator (a) the convergence curves of 240k dataset, (b) the evolution of individual loss components of 240k dataset, (c) the convergence behaviour across all three dataset scales.

with the target requirements. The improvement is most noticeable for Sym laminates, which benefit from their broader admissible orientation range and richer representation in the dataset. In contrast, DD laminates consistently show higher average errors than Quad and Sym, reflecting the limited flexibility of their four-ply repeating structure and the smaller feasible design space defined by their grammar. In terms of diversity (Fig. 10(b)), the average pairwise distance and uniqueness ratio exhibit modest variations rather than a consistent downward trend. Quad laminates show a slight rise followed by a reduction, reflecting the limited combinatorial possibilities within the $[0^\circ, \pm 45^\circ, 90^\circ]$ angle set. Sym laminates maintain the highest diversity across all dataset scales, whereas DD laminates remain relatively stable. These results together indicate that larger datasets enhance requirement alignment while preserving reasonable diversity, suggesting that the generator achieves a balanced trade-off between validity and variability. It is also worth emphasising that the lower diversity observed for the DD family in Fig. 10 is consistent with the intended role of DD laminates. A DD laminate is constructed by repeating a four-ply unit defined by only two independent angles, so its through-thickness design degrees of freedom are deliberately reduced compared with the Quad and Sym families. As a result, many DD layups lie in a relatively narrow region of the design space and exhibit very similar stiffness and performance. In this sense, the smaller spread and lower diversity of DD sequences should be interpreted as a desirable property that reflects a more constrained and easier-to-search design space, rather than as a drawback of the family or of the generative model.

To further investigate performance sensitivity, Fig. 11 presents error maps of generation accuracy across a range of P-K target combinations. The results show that regions corresponding to higher-performance targets exhibit lower generation error. Conversely, higher errors appear in low-performance regions, particularly for DD laminates which show the widest error variation. Sym laminates achieve the lowest overall error, consistent with their broader orientation space and higher achievable performance, while Quad laminates fall between the two.

Representative layups generated by the trained LayupFormer are visualised in Fig. 12, where 100 samples per laminate type are shown under fixed design targets. All generated sequences satisfy the syntactic grammar constraints, confirming the effectiveness of the domain-specific language in enforcing manufacturable designs. Distinct orientation patterns emerge for each laminate family. The DD type exhibits strongly periodic structures owing to its four-ply repeating unit, while the Quad type presents moderate variability constrained within the $[0^\circ, \pm 45^\circ, 90^\circ]$ angle set. In contrast, the Sym type demonstrates the broadest diversity, reflecting its wider admissible orientation space and higher representational freedom.

This behaviour can be explained by the interplay between the laminate grammar and the conditioning mechanism of the generator. In DD laminates, the grammar enforces strict repetition that narrows the feasible design space, leading to highly structured and less variable outputs. Quad laminates, though less constrained, still operate within a limited discrete angle set, which restricts combinatorial possibilities. Sym laminates, on the other hand, allow continuous pairing of $\pm\theta$ orientations, enabling the model to explore a broader subset of the performance-consistent design manifold. These patterns indicate that LayupFormer successfully internalises the grammar-based structural logic while maintaining flexibility to generate multiple feasible solutions under the same target conditions.

3.3. Attention-based interpretability

The attention mechanism in LayupFormer provides a direct means of examining how the model distributes focus across ply positions and conditioning inputs. Such analysis is particularly relevant for composite laminates, where mechanical coupling often depends on through-thickness interactions and mid-plane symmetry. By analysing the final-layer attention maps of both the predictor and the generator, we

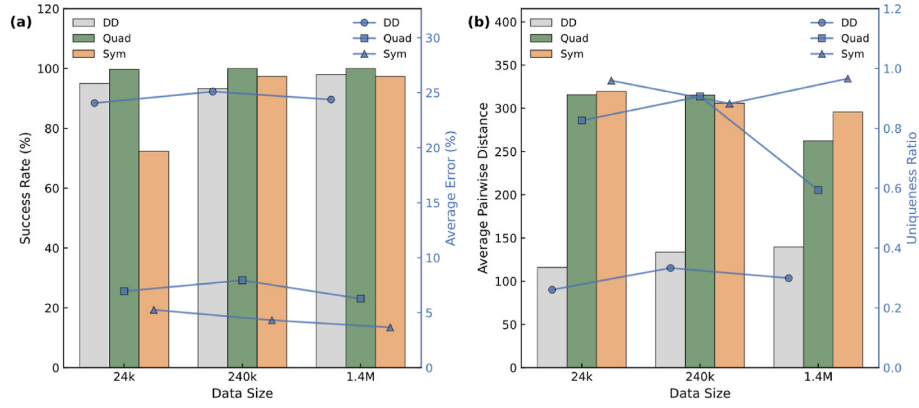


Fig. 10. Generation quality under different data scales (a) generation performance, and (b) generation diversity.

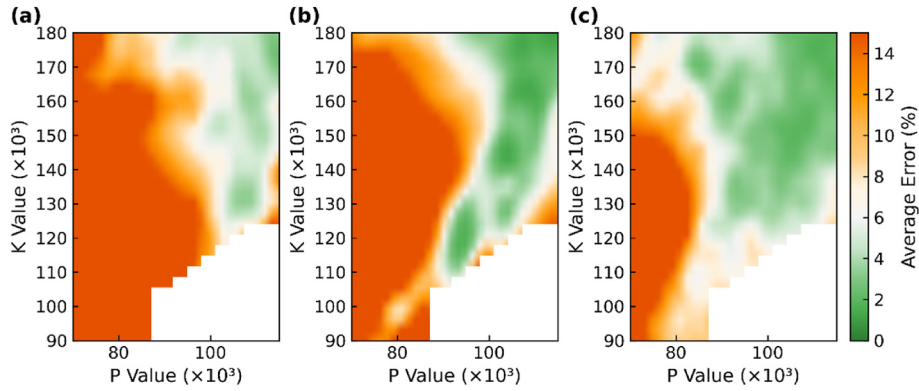


Fig. 11. Average relative-error maps, (a) DD type, (b) Quad type, (c) Sym type.

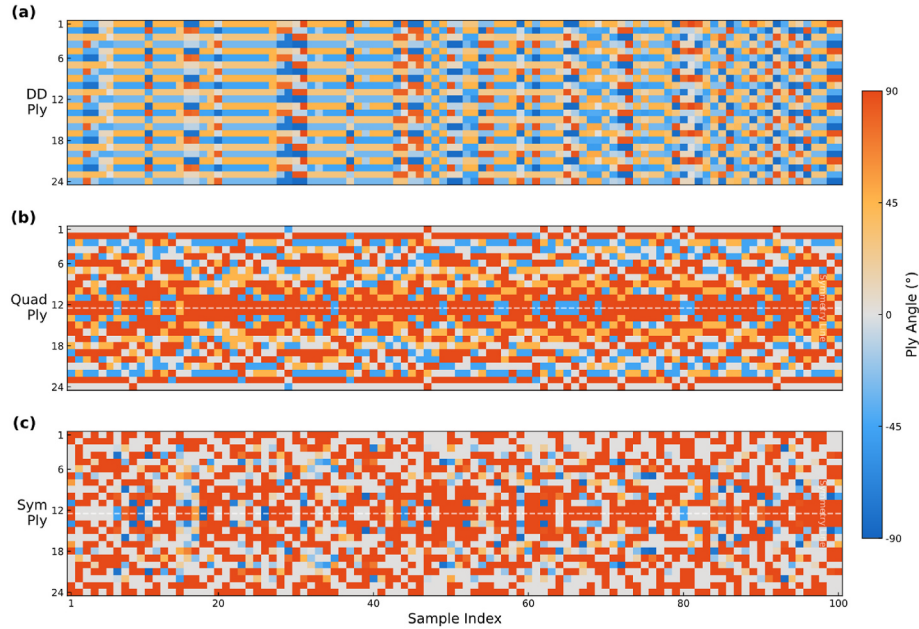


Fig. 12. Heatmap representation of 100 layups generated by LayupFormer, (a) DD type, (b) Quad type, (c) Sym type.

can gain insight into how grammar rules and performance requirements are internalised during learning and inference.

Fig. 13 presents the token-to-token attention weights of the predictor for the three laminate types. Despite the overall sparsity of attention values, several distinctive patterns are evident.

First, the attention is not limited to local neighbourhoods but extends across distant plies, indicating that the predictor has learned non-local interactions. Second, the maps reveal a clear symmetry about the laminate mid-plane, where positions t and $25-t$ show preferential attention to each other. Third, the attention structures are type

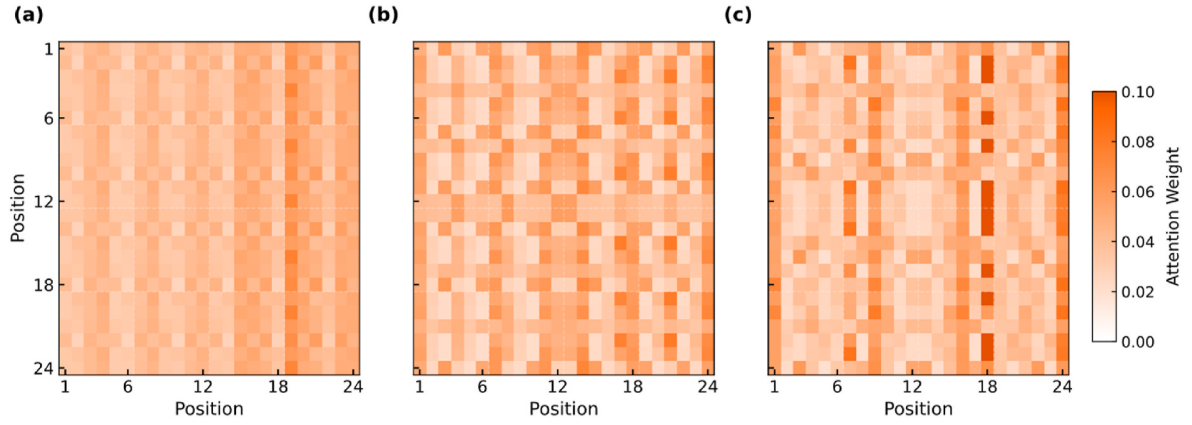


Fig. 13. Attention-weight distributions of the LayupFormer predictor model's final layer: (a) DD type, (b) Quad type, (c) Sym type.

dependent. The DD family exhibits quasi-periodic banding that reflects its four-ply repeating unit. The Quad and Sym families instead show broader and more continuous attention corridors, consistent with the mirroring rule applied to their half-stacks. These observations confirm that the predictor does not simply memorise training data but internalises meaningful structural regularities that align with the physics of laminate mechanics.

Fig. 14 illustrates the attention patterns of the generator, where Fig. 14(a) shows the masked self-attention within the decoder. The attention matrices concentrate along the principal diagonal, reflecting the autoregressive constraint. The degree of locality differs across laminate types. DD displays the sharpest diagonal concentration, whereas Quad and Sym show broader horizons, consistent with their larger orientation search spaces. Fig. 14(b) shows the cross-attention from the decoder to the conditioning memory. At the beginning of generation, the type channel dominates, providing the grammatical scaffold for sequence construction. As decoding progresses, attention gradually shifts toward the performance channel. The rate of this

transition varies across laminate types. For Sym laminates the growth of attention to performance requirements is the steepest, Quad follows with a moderate increase, while DD retains strong dependence on type information throughout.

Together, these attention analyses demonstrate that LayupFormer not only achieves accurate prediction and generation but also learns interpretable representations that align with laminate physics and grammar constraints. This interpretability provides mechanistic insight into how design rules are internalised, distinguishing the framework from conventional surrogate or optimisation-based approaches, where the design process typically remains unclear.

3.4. Comparison of DD and quad laminate families

To assess how LayupFormer exploits different laminate families in practical design, we compared the behaviour of DD and Quad laminates under the same performance requirements. Quad laminates are widely used in current engineering practice, whereas DD laminates have

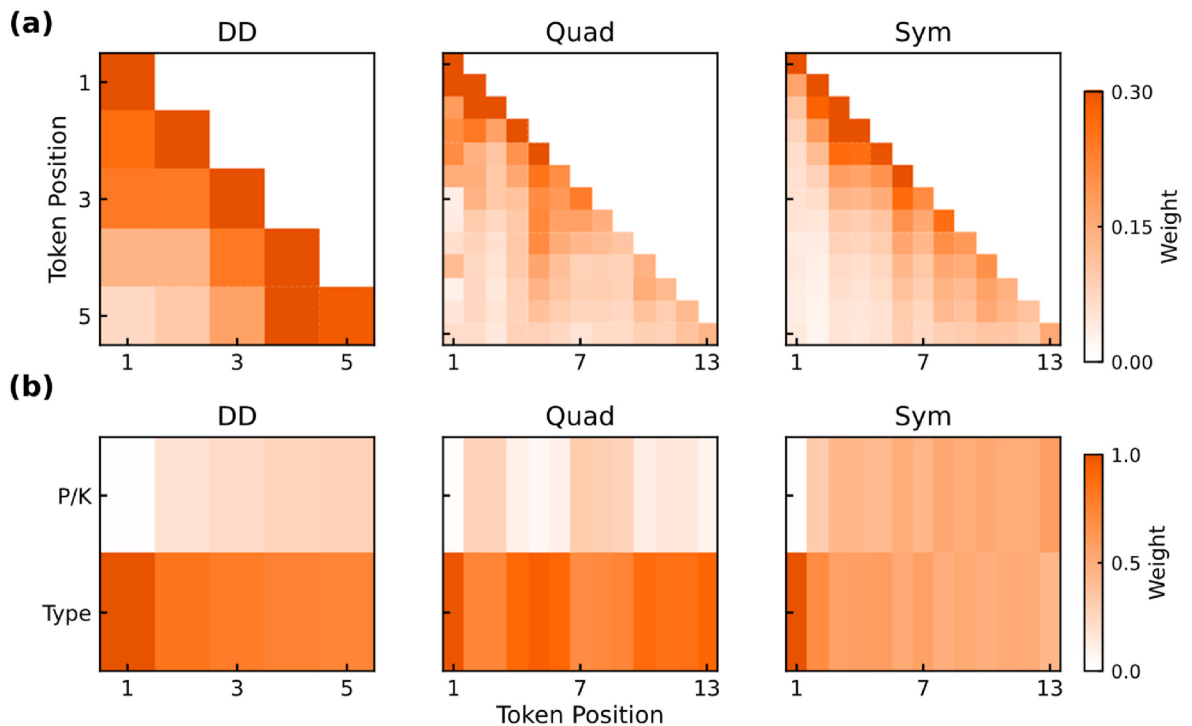


Fig. 14. Attention-weight distributions of the LayupFormer generator's final layer: (a) masked self-attention during autoregressive decoding, (b) cross-attention to the conditioning memory.

recently been proposed as a more homogeneous and simplified alternative with strong potential for automated design. A data-driven comparison between the two families is therefore essential.

We first examined a relatively demanding target, with $P = 110k$ and $K = 170k$. For this target, we performed 10,000 generation trials and collected all successful Quad and DD layups. The fibre angle frequencies over all plies were computed for each family, and five representative layups were selected by uniform sampling from the successful set, as shown in Fig. 15.

The Quad family exhibits a clear tendency to place 0° and 90° plies near the outer surfaces and close to the midplane to increase overall bearing performance, while the representative sequences remain highly diverse in terms of stacking order. In contrast, DD laminates show angles concentrated around approximately $\pm 15^\circ$, and the representative layups are formed by nearly repeated DD blocks over samples, indicating a much stronger uniqueness of the DD solutions compared with Quad for the same target.

We then extended the comparison to feasible designs over the current P–K range. A grid of target pairs (P, K) was defined, and, for each grid point, five independent generation trials were carried out for both Quad and DD to estimate the one-shot success rate of LayupFormer. The resulting normalised success maps are presented in Fig. 16. Overall, DD covers a somewhat wider region of the P–K design domain and exhibits a narrow but compact feasible band in the high stiffness and high bearing corner, whereas Quad provides feasible solutions along a diagonal band at medium to high P and K with generally higher local success rates.

Taken together, these results show that Quad offers many alternative stacking sequences in the mainstream design range, while DD produces a more concentrated set of blocks like solutions that are naturally suited to high-performance targets in the P–K space.

3.5. Experimental validation

The experimental validation was carried out using the Zoltek PX35/epoxy material system introduced in Section 2.3.4. Two laminate configurations were prepared. The first corresponds to an empirical Quad baseline widely adopted in practice, and the second is a Quad sequence proposed by LayupFormer as a candidate design with superior predicted bearing performance. The two stacking sequences are listed in Table 2.

The manufactured panels were cut into coupons following the ASTM D5961M geometry, and three specimens were tested for each configuration. The tests were carried out with a pin-loaded fixture, and the load-

displacement responses were recorded. Fig. 17 compares the nominal bearing stress-displacement curves of the two designs.

The results show that the initial bearing stiffness of the proposed layup is comparable to that of the empirical baseline. However, the LayupFormer-generated layup achieves a significantly higher mean maximum nominal bearing stress. It also displays a narrower error band, indicating improved repeatability and robustness of performance. In the post-peak region, the LayupFormer design maintains a higher load-bearing level over a larger displacement range, while the empirical baseline exhibits a more pronounced stress drop after reaching peak load.

These results provide strong evidence that the LayupFormer framework can generate laminates that outperform existing empirical designs not only in simulation but also in real-world testing. The successful experimental validation confirms the practical feasibility of the proposed approach and underscores its potential to accelerate laminate design in engineering practice.

4. Discussion

4.1. Physics-informed embedding and sample efficiency

The physics-informed embedding is a key component that enables LayupFormer to achieve high accuracy with limited data. By incorporating laminate parameters derived from Tsai's invariants into the predictor's input space, the framework introduces an inductive bias that aligns with classical laminate theory. Empirical results confirm two main outcomes. First, under limited-data conditions, the physics-informed representation achieves lower regression error than angle-only and sine-cosine encodings. Second, once trained on the full FEM dataset, the predictor reaches very high accuracy for both P and K, providing a reliable surrogate for generator training. These findings indicate that encoding periodicity and through-thickness coupling directly at the input level reduces the burden on the network to learn such structure from data, thereby improving sample efficiency without sacrificing asymptotic performance. This observation is consistent with recent studies employing physically inspired encodings for composites, where embedding laminate stiffness invariants or trace-based features has been shown to enhance model generalisation [10,24]. Compared with classical surrogate models relying purely on data-driven regression [4,23], LayupFormer's embedding strategy introduces an explicit physics link to laminate mechanics, reducing reliance on large datasets

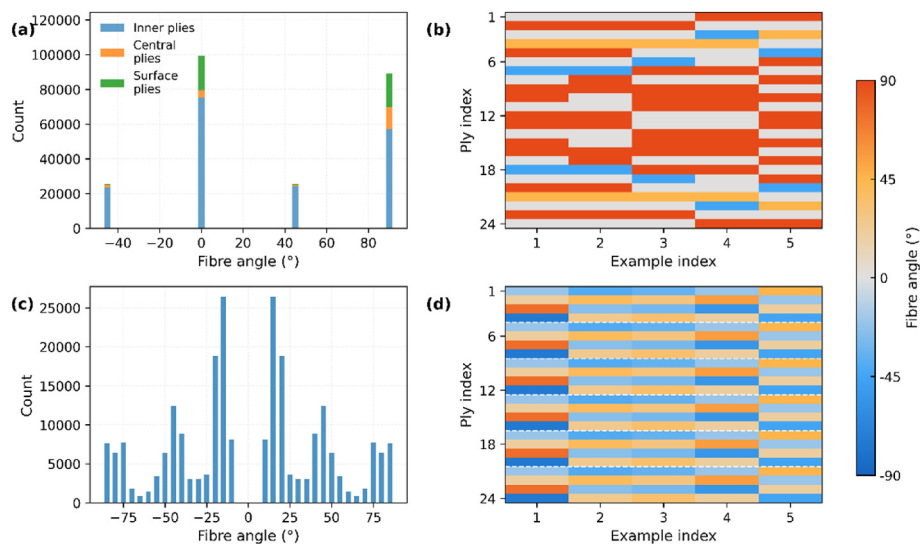


Fig. 15. Angle analysis and examples of DD and Quad laminates under a certain P–K condition: (a) Quad angle counts, (b) Generated Quad examples, (c) DD angle counts, (d) Generated DD angle counts.

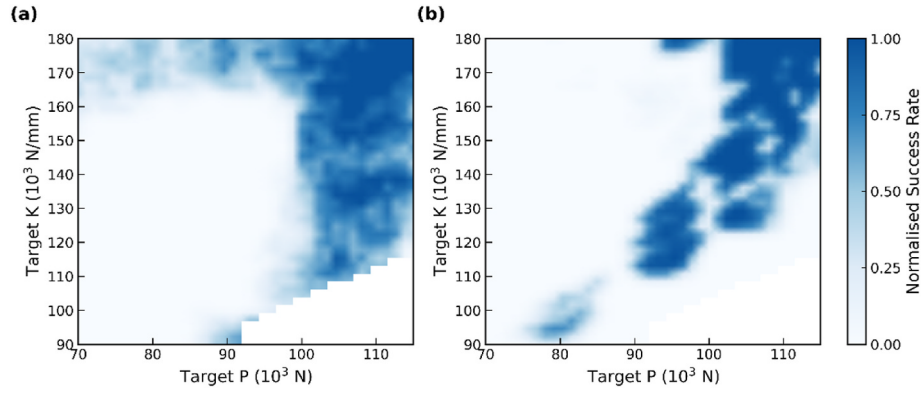


Fig. 16. The one-shot generation success maps: (a) DD family, (b) Quad family.

Table 2

The layups manufactured.

Name	Configuration
Proposed Quad Layup	[0°/0°/-45°/45°/0°/-45°/45°/0°/-45°/90°/0°/45°]s
Empirical Quad Layup	[0°/45°/-45°/90°]3s

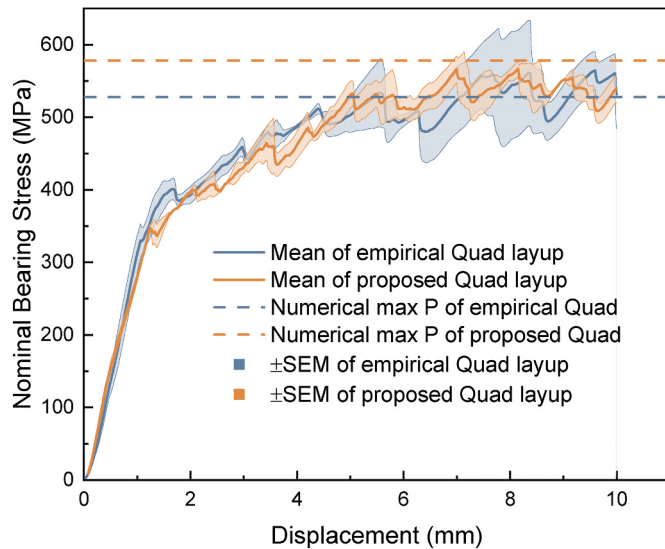


Fig. 17. Nominal bearing stress comparison.

and improving extrapolation beyond the training manifold. Such physics-informed alignment effectively bridges analytical laminate theory and modern sequence learning architectures. During generator training, coupling the decoder with a frozen predictor through a physics-consistency term further stabilises optimisation and constrains decoding to feasible regions of the design space. In this way, the embedding not only improves data efficiency but also strengthens robustness across scales.

In addition, the predictor architecture explicitly factorises material and layup information, so that the laminate response is learned as a combination of material-level features and laminate-level invariants. This structure is consistent with Tsai's scaling-law view, in which a laminate rating and a material rating contribute multiplicatively to the overall performance. In principle, once the laminate embedding has been learned from one material system, only the material branch needs to be adapted when transferring to a new material, which may further improve data efficiency for cross-material studies.

4.2. Data scale, validity, alignment, and diversity

The scale of training data governs not only convergence stability but also the trade-off between syntactic validity, requirement alignment, and design diversity. Training across three dataset scales produced stable convergence to comparable final composite losses, confirming that the training protocol is robust. However, generation quality does not plateau uniformly. Validity increases rapidly and largely saturates when trained on the medium-scale dataset. Alignment with the target performance continues to improve with additional data. At the same time, diversity metrics decline as dataset size grows. This effect is most evident for Quad laminates, whose angle set is restricted to 0°, ±45°, and 90°, and least for Sym laminates, which have the broadest admissible orientation range. The observed diversity-performance tension is a common characteristic of likelihood-based generative models. As the likelihood sharpens around high-performing and grammar-compliant regions, sampling tends to concentrate, reducing diversity. Furthermore, grammar masking and sequence expansion narrow the effective search space. These patterns highlight the inherent balance between alignment and diversity in sequence generation and emphasise the importance of data-efficient training strategies. The observed balance between alignment and diversity also reflects broader trends in generative materials design. In previous physics-constrained GAN or variational autoencoder approaches, increasing dataset size improved fidelity but often led to mode collapse and reduced structural variability [30, 31]. In contrast, the grammar-guided decoding and physics consistency in LayupFormer act as implicit regularisers that preserve syntactic validity while maintaining a balanced level of diversity. This structured trade-off represents a key advancement over stochastic sampling schemes used in earlier generative frameworks for composite design.

4.3. Attention-based interpretability

Interpretability is a distinctive advantage of LayupFormer, enabled by the analysis of attention maps in both the predictor and the generator. Attention maps provide a mechanistic view of how the framework internalises grammar and performance requirements. In the predictor, final-layer self-attention reveals long-range interactions beyond adjacent plies and symmetry about the mid-plane, with clear type-specific signatures. The DD type shows quasi-periodic banding that reflects its repeating unit, whereas the Quad and Sym types exhibit corridor-like attention structures consistent with half-stack mirroring. In the generator, masked self-attention concentrates near the diagonal under causal constraints. The degree of locality varies with laminate type, being sharpest for DD and broader for Quad and Sym. Cross-attention evolves from early dominance of the type of channel, which establishes the grammatical scaffold, toward increasing emphasis on the performance channel as decoding progresses. This transition is steepest for Sym laminates, moderate for Quad, and most persistent for DD. Together,

these analyses confirm that the model internalises physical and grammatical rules during training. The resulting interpretability renders the design process more transparent than conventional surrogate or optimisation pipelines, which typically function as black boxes. Compared with alternative interpretability routes such as graph neural networks that capture inter-ply coupling implicitly through node connections [39] and symbolic-regression-aided laminate models [32], LayupFormer provides a more direct mechanism for physical insight. The attention distributions reveal symmetry and long-range dependencies consistent with laminate mechanics, making it possible to trace how design grammar and performance requirements are jointly internalised during training. This transparency is seldom attainable in optimisation-based or adversarial generative pipelines.

4.4. Design implications of quad and DD laminate families

The results in Section 3.4 provide a data-driven perspective on how LayupFormer differentiates between Quad and DD families under identical pin-loaded bearing requirements. Quad laminates remain the industry standard for many joint designs, whereas DD laminates have been proposed as a more homogeneous alternative with strong potential for tapering, homogenisation and automation. For a high-performance target, the angle statistics in Fig. 15 show distinct design preferences. Quad solutions show 0° and 90° plies concentrated near the outer surfaces and the mid-plane and $\pm 45^\circ$ plies more frequent in the inner plies. This is consistent with classical practice for bearing and joint stiffness design and with observations that such Quad configurations may exhibit higher local strain concentrations when matched to DD laminates of similar stiffness. In contrast, DD samples generated by LayupFormer do not contain 0° ply and converge towards shallow off-axis angles, predominantly around $\pm 15^\circ$, arranged as repeated four-ply blocks. This behaviour aligns with the DD concept, in which shallow angles are used to retain most of the axial stiffness while smoothing stress and strain fields around discontinuities [40]. The model thus appears to have learned a similar trade-off between axial capacity and stress concentration from the bearing response data.

The feasible (P, K) maps in Fig. 16 highlight the complementary roles of the two families in the present 24-ply bearing case. Quad laminates form a relatively broad feasible region across the medium to high P–K range, and many distinct Quad stacking sequences achieve essentially the same target response, reflecting the strong sequence-level non-uniqueness of this discrete family. DD laminates, by contrast, occupy a narrower but compact feasible region towards the high-stiffness and high-bearing corner, with generated layups collapsing to a small number of repeated block patterns. This field-like character is consistent with experimental studies in which DD laminates have shown competitive or improved performance in unnotched compression and open-hole compression and comparable open-hole tension performance when matched to Quad laminates in stiffness [40].

The absolute success probabilities should be viewed in the context of the training data. In principle, the DD grammar is attractive for generation because it reduces the design space to repeated four-ply blocks. In practice, the empirical success rate of DD is lower than that of Quad over much of the sampled (P, K) grid, largely because the number of DD samples in the training set is much smaller and they occupy a narrower region of the design space. Despite this imbalance, LayupFormer still produces stable, highly structured DD solutions with shallow off-axis angles close to those advocated in the DD literature, rather than collapsing to arbitrary sequences. Taken together, these observations suggest that Quad acts as a robust workhorse for mainstream bearing targets within a familiar $0/\pm 45/90^\circ$ design space, while DD provides a compact set of block-based solutions that are naturally aligned with high-performance bearing targets and compatible with future extensions to thin-ply and variable-thickness DD laminates. From a manufacturing perspective, the block-based nature of DD also aligns naturally with thin-ply technologies. Commercial thin plies can be pre-ply into DD

building blocks and then deposited as single units along one axis by automated tape-laying, without the need for cross-plying. In the LayupFormer grammar, such pre-ply DD units would simply appear as single DD tokens with an effective thickness, so the same decoding strategy could in principle be reused for thin-ply DD laminates once suitable datasets become available. This supports the view that DD is particularly well suited to highly automated, thin-ply-based layup processes, whereas Quad and Sym do not admit an equally straightforward block-based implementation.

4.5. Validation scope and experimental link

The emphasis of this work lies in the LayupFormer framework, while the FEM model serves as a consistent validation mechanism rather than a research focus. The present simulations employed a linear elastic constitutive model with a Hashin-type damage criterion and stiffness degradation to extract maximum load under a fixed geometry. This configuration ensures consistency between simulation and experiment, since the latter likewise reports maximum nominal bearing load and stress. At the same time, the framework is agnostic to specific damage models. Substituting more advanced simulations would not require modifying the architecture of LayupFormer. Furthermore, the experimental consistency observed in this study aligns with analytical findings in DD laminate design [5,11], which demonstrated that repeating balanced $\pm\theta$ units enhances bearing capacity and damage tolerance. While earlier works achieved such improvements through analytical parameterisation, LayupFormer reproduces comparable structural benefits through purely data-driven inverse generation, confirming its ability to internalise fundamental laminate design principles, thereby validating the underlying physical reasoning of the framework.

Although the reported results focus on 24-ply laminates under a pin-loaded open-hole scenario with performance defined by P and K, the framework is not confined to this case. Any design objectives that can be represented as vectors may serve as conditioning inputs. Variable ply counts and richer orientation sets can be handled by extending the grammar rules introduced in Section 2.1. Likewise, alternative load cases or damage criteria can be adopted by generating the corresponding datasets. LayupFormer should therefore be regarded not as a task-specific solver but as a general paradigm for physics-informed, grammar-guided inverse laminate design.

4.6. Threats to validity

Several factors limit the scope and generalisability of the present study. First, the Hashin-based simulation may not fully capture progressive damage mechanisms such as delamination, fibre kinking, or bearing-bypass interactions. Second, the experimental validation is limited to a single laminate family, one loading case, and a small number of specimens, with simulation and experiment linked through a single scalar metric. Other standard tests used to characterise laminates, such as unnotched tension and compression, open-hole tension and compression or bending, emphasise different failure mechanisms and could alter the relative performance of different laminate families.

Furthermore, the current grammar targets balanced and quasi-symmetric families with discrete angle sets. More complex layups, including unsymmetric, tapered, or blending-constrained configurations, remain to be addressed. Within this DSL we explicitly represented Quad, DD and a six-double symmetric family, but did not introduce a separate “DD-Sym” family in which the DD block is mirrored about the mid-plane. Preliminary numerical checks for the present 24-ply bearing configuration indicated that DD and DD-Sym blocks with identical angle patterns lead to very similar bearing strength, so DD was used as the representative field-based family. This choice may not hold for other thicknesses, variable-thickness laminates or other load cases, where the distinction between DD and DD-Sym could become more relevant.

The present study focuses on a single-region, constant-thickness

laminate with 24 plies. This setup serves as a proof of concept for recasting laminate layup design as a conditional sequence generation problem. The choice deliberately simplifies the structural configuration and establishes a foundational methodology but does not capture the full complexity of realistic components such as wing skins, which are typically tapered and partitioned into many regions with varying numbers of plies. Directly applying the current high-fidelity data generation strategy to large-scale structures with multiple regions and plies would be computationally prohibitive. In addition, the present dataset and experiments are restricted to a conventional tape thickness; thin-ply laminates and pre-ply DD building blocks, which are central to many current DD manufacturing concepts, are not considered explicitly. While the block-based DD grammar is compatible with such thin-ply implementations, dedicated thin-ply datasets and experiments will be required before quantitative conclusions can be drawn for these systems.

Finally, the training data are imbalanced across laminate families and load levels. Quad laminates are much more numerous and cover a wider region of the P–K space, whereas DD samples are fewer and concentrated in a narrower band. The empirical success rates and feasible-domain comparisons reported for Quad and DD, therefore, reflect both the intrinsic structure of the laminate families and this data imbalance. Although the proposed workflow mitigates the cost of data generation by using a high-accuracy predictor trained on limited FEM data to generate large synthetic datasets, further work is needed to combine low-fidelity pre-training with high-fidelity fine-tuning and to systematically test these ideas on large, tapered, multi-region structures. These limitations underline the need for additional validation and data before the present observations can be used as general design rules.

5. Conclusion

This work presented LayupFormer, a physics-informed Transformer framework that reformulates composite laminate layup design as a requirement-conditioned sequence generation problem. By viewing layup design through the lens of a DSL, the framework directly maps target performance indicators to feasible and manufacturable stacking sequences, bridging the gap between optimisation-based search and interpretable inverse design.

The key novelty of this study lies in the integration of mechanics-based knowledge and generative modelling within a unified Transformer architecture. Specifically, the predictor introduces a Tsai-invariant embedding that grounds learning in laminate physics, while the generator employs a grammar-constrained decoding mechanism to ensure manufacturability and syntactic validity. This coupling of physics consistency, grammatical feasibility, and attention-based interpretability establishes a new paradigm for physics-informed inverse composite design.

The methodology was validated on an open-hole tensile case with 24 plies. The predictor achieved $R^2 = 0.99$ for both P and K, demonstrating highly accurate regression performance. By serving as a reliable surrogate for FEM, it substantially reduces the computational cost of data generation and enables efficient expansion of the training dataset for the generator. The generator demonstrated stable convergence across dataset scales, with validity saturating at medium scale and alignment continuing to improve with larger data. A trade-off between alignment and diversity was observed, consistent with the behaviour of likelihood-based generative models. Attention analyses further confirmed that LayupFormer internalises laminate rules and performance requirements in a physically interpretable manner. Experimental validation showed that LayupFormer-designed laminates outperform an empirical baseline in bearing tests, underscoring both the practical feasibility and predictive reliability of the framework.

Taken together, these contributions establish LayupFormer as a novel, general and extensible paradigm for physics-informed, grammar-guided laminate design. Future work will extend the framework to multi-objective conditioning, richer grammar rules, and alternative

loading scenarios. We also plan to explore interpretability-guided constraints that leverage attention patterns as implicit regularisers to enhance physical consistency and stabilise training, aiming toward a comprehensive platform for interactive and optimal composite design.

CRedit authorship contribution statement

Wenjie Xie: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Conceptualization. **Hasan Caglar:** Validation, Formal analysis, Data curation. **Kailun Deng:** Software, Methodology, Data curation. **David Ayre:** Supervision, Resources, Funding acquisition. **Yifan Zhao:** Writing – review & editing, Supervision, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability

Data will be made available on request.

References

- [1] B. Vermes, S.W. Tsai, A. Riccio, et al., Application of the Tsai's modulus and double-double concepts to a new affordable design approach for composite laminates, *Compos. Struct.* 259 (2021) 113246.
- [2] B.L. Ennis, et al., Economic competitiveness of pultruded fiber composites for wind turbine applications, *Composites, Part B* 265 (2023) 110960.
- [3] L. Jia, C. Zhang, J. Li, L. Yao, C. Tang, Validation and development of trace-based approach for composite laminates, *Compos. Sci. Technol.* 221 (2022) 109348.
- [4] A. Şerban, Failure estimation of the composite laminates in layup optimization using finite element analysis and deep learning, *J. Mater. Eng. Perform.* 29 (2020) 6582–6594.
- [5] E. Kappel, Double-double laminates for aerospace applications—Finding best laminates for given load sets, *Composites Part C: Open Access* 8 (2022) 100244.
- [6] A.T. Nettles, *Basic Mechanics of Laminated Composite Plates*, vol. 1351, NASA Reference Publication, 1994. NASA Technical Report.
- [7] S.W. Tsai, J.D.D. Melo, An invariant-based theory of composites, *Compos. Sci. Technol.* 100 (2014) 237–243.
- [8] S.W. Tsai, J.D.D. Melo, *Composite Materials Design and Testing: Unlocking Mystery with Invariants*, Composites Design Group, Stanford, UK, 2015.
- [9] S.W. Tsai, S. Sih, J.D.D. Melo, Trace-based stiffness for a universal design of carbon-fiber reinforced composite structures, *Compos. Sci. Technol.* 118 (2015) 23–30.
- [10] L.L. Vignoli, J. Gomide, L.E. Santana, Y.S. Macedo, J.S. Oliveira, Using asymptotic homogenization and machine learning to derive a 3D trace theory for GFRP, *Mech. Adv. Mater. Struct.* 31 (30) (2024) 12665–12677.
- [11] S.W. Tsai, Double-double: new family of composite laminates, *AIAA J.* 59 (11) (2021) 4293–4305.
- [12] R.M. Koide, G.V.Z.D. França, M.A. Luersen, An ant colony algorithm applied to layup optimization of laminated composite plates, *Lat. Am. J. Solid. Struct.* 10 (2013) 491–504.
- [13] F. Aymerich, M. Serra, Optimization of laminate stacking sequence for maximum buckling load using the ant colony optimization (ACO) metaheuristic, *Composites Part A* 39 (2008) 262–272.
- [14] M. Abachizadeh, M. Tahani, An ant colony optimization approach to multi-objective optimal design of symmetric hybrid laminates for maximum fundamental frequency and minimum cost, *Struct. Multidiscip. Optim.* 37 (2009) 367–376.
- [15] W. Wang, S. Guo, N. Chang, Z. Feng, W. Yang, A modified ant colony algorithm for the stacking sequence optimization of a rectangular laminate, *Struct. Multidiscip. Optim.* 41 (2010) 711–720.
- [16] W. Wang, S. Guo, N. Chang, W. Yang, Optimum buckling design of composite stiffened panels using ant colony algorithm, *Compos. Struct.* 92 (2010) 712–719.
- [17] Y. Terada, A. Todoroki, Y. Shimamura, Stacking sequence optimizations using fractal branch and bound method for laminated composites, *JSME International Journal, Series A* 44 (2001) 490–498.
- [18] M. Akbulut, F.O. Sonmez, Optimum design of composite laminates for minimum thickness, *Comput. Struct.* 86 (2008) 1974–1982.

- [19] Z. Jing, X. Fan, Q. Sun, Stacking sequence optimization of composite laminates for maximum buckling load using permutation search algorithm, *Compos. Struct.* 121 (2015) 225–236.
- [20] Z. Jing, X. Li, Q. Sun, K. Liang, Y. Zhang, L. Duan, A 2D-sampling optimization method for buckling layup design of doubly-curved laminated composite shallow shells, *Compos. Struct.* 297 (2022) 115934.
- [21] Z. Jing, L. Duan, A three-dimensional sampling optimization method for buckling optimization of variable prestressed composite plates, *Thin-Walled Struct.* 181 (2022) 110096.
- [22] J.E. Herencia, R.T. Haftka, P.M. Weaver, M.I. Friswell, Lay-up optimization of composite stiffened panels using linear approximations in lamination space, *AIAA J.* 46 (9) (2008) 2387–2391.
- [23] C. Furtado, L.F. Pereira, R.P. Tavares, M. Salgado, F. Otero, G. Catalanotti, A. Arteiro, M.A. Bessa, P.P. Camanho, A methodology to generate design allowables of composite laminates using machine learning, *Int. J. Solid Struct.* 233 (2021) 111095.
- [24] S. Khalid, M.H. Yazdani, M.M. Azad, M.U. Elahi, I. Raouf, H.S. Kim, Advancements in physics-informed neural networks for laminated composites: a comprehensive review, *Mathematics* 13 (1) (2025). Article 17.
- [25] H. Hong, W. Kim, S. Kim, K. Lee, S.S. Kim, Deep transfer learning for efficient and accurate prediction of composite pressure vessel behaviors, *Compos. Appl. Sci. Manuf.* 186 (2024) 108413.
- [26] X. Liu, C.A. Featherston, D. Kennedy, Two-level layup optimization of composite laminate using lamination parameters, *Compos. Struct.* 211 (2019) 337–350.
- [27] Z.R. Wu, Y. Yang, H. Lei, Progressive failure analysis of perforated composite laminates considering nonlinear shear effect, *Sci. Rep.* 13 (2023) 22375.
- [28] A. Choudhury, S.C. Mondal, S. Sarkar, Failure analysis of laminated composite plate under hygro-thermo mechanical load and optimisation, *Int. J. Appl. Mech. Eng.* 24 (3) (2019) 509–526.
- [29] D.W. Sleight, N.F. Knight Jr., J.T. Wang, Evaluation of a progressive failure analysis methodology for laminated composite structures, in: 38th AIAA Structures, *Structural Dynamics and Materials Conference*, AIAA, 1997. Paper 97-1187.
- [30] Y. Zhang, J. Gu, Z. Li, S. Ruan, C. Shen, Data-driven lay-up design of a type IV hydrogen storage vessel based on physics-constrained generative adversarial networks (PCGANs), *J. Energy Storage* 98 (2024) 113130.
- [31] C. Qiu, Y. Han, L. Shanmugam, Y. Zhao, S. Dong, S. Du, J. Yang, A deep learning-based composite design strategy for efficient selection of material and layup sequences from a given database, *Compos. Sci. Technol.* 230 (2022) 109154.
- [32] C. Qiu, H. Song, J. Yang, A data-driven approach to identify the optimal sub-laminates for homogeneity design under the concept of double-double composites, *Compos. Appl. Sci. Manuf.* 195 (2025) 108897.
- [33] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need, *Adv. Neural Inf. Process. Syst.* 30 (2017).
- [34] J. Jumper, R. Evans, A. Pritzel, et al., Highly accurate protein structure prediction with AlphaFold, *Nature* 596 (2021) 583–589.
- [35] C. Xu, Y. Wang, A. Barati Farimani, TransPolymer: a transformer based language model for polymer property predictions, *npj Comput. Mater.* 9 (1) (2023). Article 64.
- [36] Y. Etesam, H. Cheong, M.M. Ataei, P.K. Jayaraman, Deep generative model for mechanical system configuration design, *Proc. AAAI Conf. Artif. Intell.* 39 (16) (2025) 16496–16504.
- [37] Z. Hashin, Fatigue Failure Criteria for Unidirectional Fiber Composites, 1980. No. TR5).
- [38] ASTM International, Standard Test Method for Bearing Response of Polymer Matrix Composite Laminates, ASTM International, West Conshohocken, PA, 2023. ASTM D5961/D5961M-23).
- [39] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, P.S. Yu, A comprehensive survey on graph neural networks, *IEEE Transact. Neural Networks Learn. Syst.* 32 (1) (2020) 4–24.
- [40] A. Hawkins, S.L. Millen, A. Aravand, Double-double laminates: a comprehensive assessment of the mechanical performance in comparison to traditional quad laminates, *Compos. B Eng.* (2025) 113115.



DSPACE

<https://dspace.org/>

LayupFormer: a deep generative model for composite laminate layup design

Xie, Wenjie; Caglar, Hasan; Deng, Kailun; Ayre, David; Zhao, Yifan

2026-03-01

Elsevier

<https://doi.org/10.1016/j.compscitech.2025.111490>

Xie W, Caglar H, Deng K, et al., (2026) LayupFormer: a deep generative model for composite laminate layup design. Composites Science and Technology, Volume 275, March 2026, Article number 111490