Bayesian optimisation of resolved interlaminar stresses in composite laminates using high-aspect-ratio discontinuous Galerkin finite elements

Soheil Navvabi[1*], Stefano Giani[1] and William M. Coombs[1]

[1*]Department of Engineering, Durham University, Durham, DH1 3LE, UK.

*Corresponding author(s). E-mail(s): soheil.navvabi@durham.ac.uk;

Abstract

Laminate optimisation is a key route to improved composite structures, linking material choice, fibre orientation and stacking sequence to structural performance and manufacturability. Using this design freedom to optimise local interlaminar-stress severity directly is much harder, since the governing stress field is three-dimensional, highly localised and costly to resolve accurately. This paper presents a coupled discontinuous Galerkin (DG) finite-element and Bayesian optimisation (BO) framework for treating local interlaminar-stress severity as the optimisation objective itself. Candidate laminates are chosen from a finite material-and-fibre-angle library and evaluated with a full 3D DG finite-element solver, whose high-aspect-ratio meshes are central to the efficiency of the optimisation framework: they concentrate refinement around monitored stress regions and resolve the local three-dimensional stress fields without requiring fine elements throughout the surrounding domain. Scalar patch objectives are constructed from the resolved interlaminar normal and shear stresses, including hardmax, smoothmax and monotonic power/log transforms, to test how objective representation affects BO performance without changing the underlying candidate laminate set. Benchmarks are straight free-edge and corner-stressed rectangular-cutout laminates. Against a matched genetic-algorithm (GA) baseline, BO consistently gives higher exact-optimum reliability and better final ranks. In the free-edge benchmark, the best transformed objectives identify the optimum in every run; GA needs 7 times more evaluations and 6.7 times longer mean time to identify the same optimum. In the cutout benchmark, BO finds the optimum after sampling only about 1.2% of the feasible space; GA needs 8 times more evaluations and 7.5 times longer mean time to do so.

Keywords: Composite laminates, Laminate optimisation, Interlaminar stress, Bayesian optimisation, High-fidelity optimisation, Discontinuous Galerkin

Article Highlights

- High-aspect-ratio discontinuous Galerkin meshes make local interlaminar-stress optimisation practical.
- Bayesian optimisation finds exact optima with up to 8 times fewer laminate evaluations than genetic algorithms.
- Bayesian optimisation found the cutout optimum in 10/10 runs after sampling only 1.2% of the design space.

1 Introduction

Composite laminate optimisation is central to exploiting anisotropy rather than merely accommodating it: ply materials, fibre directions and stacking order can be chosen to tune stiffness, strength, buckling resistance, dynamic response and manufacturability within the same structural envelope. This same design freedom also makes laminate optimisation a demanding search problem. The mechanical response of a laminate is governed by both the properties of the constituent plies and the laminate architecture, including ply order, fibre orientation, thickness and material assignment through the laminate thickness. Laminate design therefore naturally leads to a discrete or mixed discrete-continuous optimisation problem in which many stacking sequences may satisfy the same design rules while producing markedly different structural responses. The difficulty of searching this stacking-sequence space was recognised early in genetic algorithm (GA) studies, where ply orientations and ply order were treated directly as optimisation variables. Later reviews show that the same challenge persists across constant-stiffness, variable-stiffness and nonconventional laminate design: the feasible space is large, non-convex and often contains many designs with similar performance [1, 8, 20, 21]. The design-variable choice therefore shapes the optimisation problem itself: lamination parameters, ply-count variables, angle sets and direct stacking-sequence descriptions each expose a different representation of the same laminate-design space [35]. Such design representations make the search space more tractable. In practice, this is often paired with reduced-order models and global response measures, which further reduce cost but limit direct access to more local response quantities, such as three-dimensional stress fields.

This compromise is particularly visible in large-scale laminate design workflows, where the full stacking space is rarely searched directly. Instead, the optimisation is often decomposed into a reduced continuous problem followed by a discrete stack-recovery or blending step [29, 39]. Lamination parameters provide one such route, replacing the ply-by-ply stacking description with stiffness-space variables that are smoother and lower-dimensional, but that must later be converted into manufacturable stacking sequences [1, 35]. The same separation between continuous target design and discrete layup recovery appears in workflows based on generic stacks, where a gradient-based sizing step defines target thickness and stiffness distributions that are then matched by mixed-integer stacking and blending optimisation to obtain manufacturable layups for large aerospace structures [39]. Stacking sequence tables and their recent extensions similarly encode ply continuity, ply drops and regional thickness changes in a way that reduces invalid designs and improves the efficiency of variable-thickness optimisation [41]. By using compact laminate representations together with response quantities that are cheap to compute, these strategies make the search computationally feasible for global or structural-level objectives such as stiffness, buckling load or weight. Their limitation appears when the objective is governed by local three-dimensional stress concentrations, where the relevant quantity cannot be captured reliably by reduced laminate or plate-level response measures. In that case, the stress field must be evaluated within the optimisation itself, rather than inferred after the stack has already been selected [23, 29, 39].

Alongside these design-space reductions, a large body of work has focused on the search algorithm itself. GA-based methods provided an early and influential route for exploring stacking-sequence spaces [8], and later studies introduced specialised deterministic and metaheuristic searches to improve scalability, diversity and convergence. For example, fractal branch-and-bound methods exploit lamination-parameter structure to prune large parts of the stacking-sequence tree, beam-search strategies retain only promising partial layups when many plies are involved, and adaptive harmony search provides another population-based alternative for discrete laminate design [14, 26, 51]. These methods are valuable because they address the combinatorial nature of stacking directly, rather than replacing ply order with a purely continuous approximation. However, their practical performance is closely tied to the cost of evaluating each candidate design. In many benchmark problems, the objective can be evaluated from analytical or reduced-order structural models. Under those conditions, thousands of candidate stacks can be searched at acceptable cost. The same strategy becomes much less attractive when each candidate requires a high-fidelity analysis, for example to resolve a three-dimensional stress field. In that case, the limiting factor shifts from the search algorithm alone to the coupled cost of search and analysis.

Where greater physical realism is needed, finite-element analysis has become a natural partner for laminate optimisation, because it can represent more realistic structural geometries, boundary conditions, load paths and failure measures than closed-form laminate calculations. In aerospace-scale applications, for example, multi-laminate wing-box optimisation has been coupled directly to Abaqus shell models, but the finite-element solution was reported to dominate the computational time and required hundreds of analysis submissions [48]. Other studies have improved tractability through strategies such as global-local submodelling, analytical sensitivities and gradient-based solvers. For example, global shell models can be combined with local shell submodels to evaluate selected responses more efficiently [42], while interior-point optimisation has been shown to reduce the number of Abaqus buckling analyses compared with population-based GA or differential-evolution searches [38]. Similarly, fatigue-driven local damage measures can be incorporated through aggregation, scaling and adjoint sensitivities in equivalent-single-layer shell optimisation [22]. Together, these developments show that once laminate optimisation is coupled to finite-element analysis, the cost of repeated structural evaluations becomes a central limitation. At the same time, the analysis models used to manage this cost are still predominantly shell, plate, equivalent-single-layer or shell-based global-local formulations. They are valuable for structural-level responses, but they do not provide the resolved through-thickness stress description needed for direct interlaminar-stress objectives. This leaves a need for an analysis model that retains three-dimensional stress resolution while remaining efficient enough for repeated optimisation calls.

One response to this cost bottleneck has been to replace repeated direct analyses with surrogate or machine-learning models. Early work on stacking-sequence optimisation used local regression response surfaces inside a GA so that only selected candidates required full structural evaluation [27]. Two-level approximation strategies extended this idea by building approximate structural responses from a limited number of finite-element and sensitivity analyses, allowing genetic search or sizing steps to proceed without calling the full solver at every candidate [4, 9]. More recent work has used machine-learning surrogates in a similar spirit: artificial neural networks trained on Abaqus buckling data have been used to replace repeated shell eigenvalue analyses during laminate weight optimisation [28], while random-forest models trained from validated impact simulations have been used to support multi-objective optimisation of nonconventional ply angles [10]. Taken together, these studies show that surrogate assistance can reduce repeated solver calls once a sufficiently informative surrogate model is available, a trend also reflected in broader reviews of machine learning in laminated composites [50]. For the present class of objectives, however, building that surrogate model is itself costly, because each training label requires a resolved three-dimensional stress analysis, and a substantial number of such labels may be needed before the model can reliably rank candidates in a large discrete stacking space.

This shifts the cost problem to the data acquisition stage: the optimiser must decide which high-fidelity analyses are worth performing. Sequential surrogate-based optimisation addresses this by updating a surrogate as new data are acquired. Bayesian optimisation (BO) gives this process an acquisition-driven form, in which the probabilistic surrogate quantifies both predicted performance and uncertainty, and the acquisition rule uses this information to choose the next high-fidelity evaluation [16]. This makes BO especially well suited to discrete or mixed laminate-design problems with expensive high-fidelity evaluations, because it can build the surrogate during the search rather than requiring a costly offline surrogate model in advance [53]. This logic has already been demonstrated for laminate stacking selection: a Gaussian-process BO framework was coupled to a three-dimensional progressive-damage finite-element model to improve open-hole tensile strength, with the acquisition process reducing the need for blind exploration of the stacking space [11]. BO has also been developed for categorical and integer-valued variables, which is directly relevant when material choices and fibre-angle libraries are treated as discrete design inputs [18]. Together, these studies establish BO as a credible tool for expensive laminate-design problems, but they also show that acquisition-driven sampling solves only one side of the cost problem. For local objectives that require full three-dimensional stress resolution, practical optimisation also needs an evaluator that is accurate enough to resolve the relevant response and efficient enough to be called repeatedly. The remaining need is therefore to pair BO's sample-efficient search with an analysis model tailored to accurate, repeated evaluation of local three-dimensional response quantities.

The need for such an analysis model becomes clearest when the optimisation objective is connected to interlaminar stress, delamination, ply drops or free-edge effects. These phenomena are local and three-dimensional by nature, so they are difficult to represent using the same reduced models that make global stack optimisation efficient. In blended laminates, ply drops are recognised as sources of out-of-plane stress concentration and possible delamination, yet optimisation frameworks based on stacking sequence tables typically manage this risk through blending and design rules rather than by computing the local interlaminar stress field [23]. Similarly, progressive-damage optimisation of composite tubes can include material degradation in a finite-element loop, but shell and plane-stress assumptions remain limiting when out-of-plane shear or delamination-dominated failure mechanisms are important [2]. Multi-scale stiffened-panel design offers another instructive example: economical shell models can drive the optimisation, while refined layer-wise analysis is used only afterwards to examine transverse stresses that were absent from the optimisation objective [33]. Experimental and numerical studies of variable-thickness or unconventional laminates reach a similar conclusion, identifying local ply-drop effects or local criteria beyond global buckling as important but too costly to resolve directly during optimisation [13, 34]. This cost barrier is also explicit in gradient-based laminate optimisation, where higher-order shells or 3D finite elements are recognised as necessary for out-of-plane failure effects but generally considered too expensive for routine optimisation [49].

Recent optimisation frameworks show how far the field has progressed toward scalable, manufacturable and industrially relevant laminate design. For aircraft wing skins, for example, two-stage workflows can first optimise continuous sizing and stiffness targets, and then use generic stacks and mixed-integer programming to convert those targets into blended ply-by-ply layups for large structures [40]. A related mixed-variable framework applies the same manufacturability-driven logic to both conventional laminate stacks and Double-Double layups, using bi-level outer approximation to select layups from catalogues while enforcing manufacturing compatibility, drop-off and variability controls for industrial panel and spar designs [17]. Variable-stiffness design has progressed along a complementary path, using level-set towpath descriptions to produce manufacturable fibre steering under automated-fibre-placement constraints, while gradient-based optimisation minimises smooth aggregations of Tsai-Hill/Tsai-Wu failure indices within a classical plate-theory framework [15]. Taken together, these examples and broader reviews of high-fidelity and metamodel-based laminate optimisation show that modern laminate design pipelines have become increasingly sophisticated in coupling search efficiency, manufacturability and structural performance [24]. The remaining gap is more specific: the response models used in these pipelines are still predominantly classical laminate theory, shell/plate, global finite-element, reserve-factor, failure-index or industrial sizing/certification analyses. They therefore advance manufacturable laminate optimisation, but do not provide a compact optimisation loop in which the objective itself is a resolved local three-dimensional interlaminar-stress field.

Taken together, the literature shows substantial progress on each part of the laminate optimisation problem. Stacking-sequence design has been studied as a discrete and mixed-variable search problem; manufacturability has been addressed through lamination parameters, generic stacks, blending rules and stack-recovery methods; finite-element models have been coupled to optimisers for more realistic structural responses; and surrogate, metamodel and Bayesian strategies have been introduced to reduce the number of expensive evaluations [11, 15, 17, 24, 40]. At the same time, the papers concerned with ply drops, out-of-plane failure and local interlaminar effects show why such stresses are difficult to include in this same loop: resolving them requires a local three-dimensional stress description, while repeated high-fidelity evaluations can make direct stack search impractical [13, 49]. The remaining need is therefore not a generic laminate optimiser, nor a generic high-fidelity analysis, but a framework in which the analysis model and the optimiser are matched to the same local stress objective. Against this background, the present work combines an efficient full 3D discontinuous Galerkin (DG) finite-element formulation [37] with BO. The DG evaluator is the analysis-side enabler: by allowing high-aspect-ratio elements, it can concentrate small, locally refined elements around regions where three-dimensional interlaminar effects are important and keep the mesh coarse elsewhere, substantially reducing the element and degree-of-freedom count while retaining accurate resolved local stress information. BO provides the sample-efficient sequential decision layer, using previous evaluations to decide which discrete laminate should receive the next high-fidelity analysis. It therefore avoids the many population-level objective queries typical of evolutionary search, as well as the large and costly upfront training datasets required by offline surrogate or machine-learning models. This DG–BO coupling allows local interlaminar-stress severity to be treated not as an a posteriori check or design-rule proxy, but as the optimisation objective itself, and it is tested in both a canonical free-edge setting and a geometry-induced cutout-corner setting.

2 Framework and methodology

This section defines the coupled analysis-optimisation formulation used to search for laminate designs with reduced local interlaminar-stress severity. Each candidate design is first interpreted as a discrete laminate configuration, specified by the material and fibre-angle choices allowed by the design library. A high-fidelity DG finite-element solver then evaluates its mechanical response, and the relevant local interlaminar stress components are extracted over a monitored region. The resulting stress fields are reduced to scalar objective values, providing the response data used by BO to select further candidate laminates under a limited evaluation budget. The two main components of the framework therefore have distinct roles: the DG evaluator provides physically resolved local stress information in regions where three-dimensional interlaminar effects are important, while BO provides the decision layer that uses the evaluated objective values to allocate subsequent high-fidelity analyses more selectively

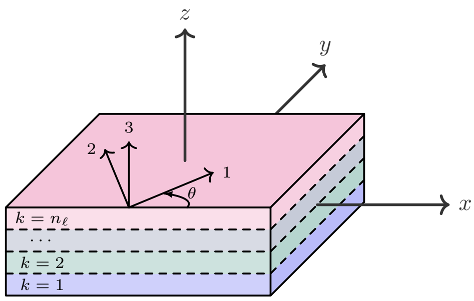


Fig. 1 Schematic representation of the laminate design parameterisation. Each independently specified layer or design group $k \in \{1, \ldots, n_\ell\}$ is assigned a material choice m_k and an in-plane fibre angle θ_k. The 1–2–3 axis marker shows the local material axes for one representative layer or design group: axis 1 is the fibre direction and axis 2 is the in-plane transverse direction, both rotated in the laminate plane by θ_k, while axis 3 remains the through-thickness direction

than an exhaustive search would allow. A GA is also defined as a population-based baseline, so that the BO results can be interpreted against a standard derivative-free search method using the same design space, objectives, and success criteria.

2.1 Laminate design space

The optimisation is carried out over a discrete laminate design space. Material choices are drawn from a finite library $\mathcal{M}$ containing n_m admissible lamina systems, and fibre-angle choices are drawn from a finite set Θ containing n_θ admissible orientations. These libraries are specified by the benchmark or application being studied. For a laminate parameterisation with n_ℓ independently specified layers or design groups, a candidate design is represented by the integer-coded vector

$$\boldsymbol{x} = (m_1, \theta_1, \ldots, m_{n_\ell}, \theta_{n_\ell}), \tag{1}$$

where $m_k \in \mathcal{M}$ denotes the material choice for layer or design group k, and $\theta_k \in \Theta$ denotes the corresponding fibre-angle choice. Each layer or design group can therefore select its own material and orientation, as illustrated in Fig. 1. This representation makes the design variables directly compatible with both BO and the GA baseline, while the decoding step converts each integer code into the orthotropic lamina properties and fibre angle required by the DG evaluator.

For these finite libraries, the unfiltered design library is denoted by $\mathcal{X}_{\mathrm{raw}}$, with

$$|\mathcal{X}_{\mathrm{raw}}| = n_m^{n_\ell} n_\theta^{n_\ell}. \tag{2}$$

The optimiser is allowed to change only the material and fibre-angle choices in Eq. (1). For each problem instance, the specimen geometry, prescribed-displacement boundary conditions, laminate thickness pattern, mesh definition, and monitored stress region are fixed. These quantities define the mechanical problem being evaluated, not additional optimisation variables. The feasible design space is written generally as

$$\mathcal{X}_{\mathrm{feas}} = \{\boldsymbol{x} \in \mathcal{X}_{\mathrm{raw}} : \boldsymbol{x} \text{ satisfies the prescribed feasibility rules}\}. \tag{3}$$

These feasibility rules are problem dependent: they specify which material-angle combinations are allowed for the chosen laminate parameterisation and benchmark or application. The design-space definition is summarised in Table 1.

Table 1 General discrete laminate design-space definition.

Item	Definition
Design vector	$\boldsymbol{x} = (m_1, \theta_1, \ldots, m_{n_\ell}, \theta_{n_\ell})$, with independent material and fibre-angle choices for each layer or design group.
Material choices	$\mathcal{M}$: finite library of n_m admissible lamina systems.
Fibre-angle choices	Θ: finite set of n_θ admissible fibre angles.
Fixed during optimisation	Specimen geometry, prescribed-displacement boundary conditions, laminate thickness pattern, mesh definition, and monitored stress region.
Feasibility rule	Problem-specific restrictions defining $\mathcal{X}_{\text{feas}} \subseteq \mathcal{X}_{\text{raw}}$.
Raw design count	$n_m^{n_\ell} n_\theta^{n_\ell}$ configurations before feasibility filtering.

2.2 DG stress-analysis evaluator

In this study, each candidate laminate is evaluated using the DG finite-element solver developed, verified, and validated for three-dimensional composite-laminate analysis in the earlier DG study [37]. The solver computes the three-dimensional linear-elastic response on the laminate domain Ω,

$$-\boldsymbol{\nabla} \cdot \boldsymbol{\sigma}(\mathbf{u}) = \mathbf{f}. \tag{4}$$

where $\mathbf{u}$ is the displacement field, $\boldsymbol{\sigma}$ is the Cauchy stress, and $\mathbf{f}$ is the body force. The stress is related to the infinitesimal strain by $\boldsymbol{\sigma}(\mathbf{u}) = \mathbf{D}(\boldsymbol{x})\, \boldsymbol{\varepsilon}(\mathbf{u})$, with $\boldsymbol{\varepsilon}(\mathbf{u}) = \frac{1}{2}(\boldsymbol{\nabla}\mathbf{u} + \boldsymbol{\nabla}\mathbf{u}^{\mathsf{T}})$ and $\mathbf{D}(\boldsymbol{x})$ the piecewise-constant orthotropic stiffness tensor defined by the candidate laminate design. In the benchmark problems considered here, the external loading is imposed through prescribed boundary displacements, so the body-force term is zero. The material constants and fibre orientations selected by a candidate design specify the stiffness tensors assigned to the laminate layers. For each layer or design group $k \in \{1, \ldots, n_\ell\}$, the orthotropic stiffness is first formed in its local material axes and then transformed into the global coordinate system using the prescribed in-plane fibre angle,

$$\mathbf{D}^{(k)}(\theta_k) = \mathbf{T}_\sigma^{-1}(\theta_k)\, \mathbf{D}_0^{(k)}\, \mathbf{T}_\varepsilon(\theta_k), \tag{5}$$

where $\mathbf{D}_0^{(k)}$ is the local orthotropic stiffness matrix. The matrices $\mathbf{T}_\sigma$ and $\mathbf{T}_\varepsilon$ are the corresponding Voigt stress and engineering-strain transformations, defined by

$$\boldsymbol{\sigma}_{\text{loc}} = \mathbf{T}_\sigma(\theta_k)\, \boldsymbol{\sigma}_{\text{glob}}, \qquad \boldsymbol{\varepsilon}_{\text{loc}} = \mathbf{T}_\varepsilon(\theta_k)\, \boldsymbol{\varepsilon}_{\text{glob}}, \tag{6}$$

with Voigt ordering $[xx, yy, zz, xy, yz, zx]$ and engineering shear strains. The displacement field is approximated on hexahedral elements using a symmetric interior penalty

DG formulation. Unlike conforming finite elements, the displacement approximation is allowed to be discontinuous between neighbouring elements. Interelement continuity and traction transfer are imposed weakly: consistent numerical fluxes transfer tractions across faces, while penalty terms constrain the displacement jumps permitted by the discontinuous approximation. The face-wise penalty is scaled as

$$\eta_E = \gamma \, \frac{p_E^2}{h_E^{\perp}} \, L_E, \tag{7}$$

where p_E is the polynomial degree on face E, $h_E^{\perp}$ is the effective element size measured normal to the face, and L_E is a stiffness measure constructed from the elasticity tensors adjacent to that face. For an interior face $E = \partial K^+ \cap \partial K^-$, this measure is $L_E = \omega_+^2 \delta_+ + \omega_-^2 \delta_-$, with $\delta_\pm = \|\mathbf{A}^\pm\|_F$, $\omega_+ = \delta_-/(\delta_+ + \delta_-)$, and $\omega_- = \delta_+/(\delta_+ + \delta_-)$, where $\mathbf{A}^\pm$ are the normal-stiffness projections of the neighbouring ply tensors $\mathbf{D}^\pm$ along the face normal. This scaling ties the penalty directly to the face-normal mesh size and the stiffness contrast across the interface, supporting stable traction transfer on stretched hexahedral elements as well as across neighbouring orthotropic plies.

The key advantage for the present optimisation problem is that the mesh can remain coarse over most of the specimen while being refined in regions where local interlaminar stress gradients are expected. In laminated structures, such gradients are typically concentrated near free edges, ply interfaces, geometric discontinuities, or monitored stress patches. A conventional low-aspect-ratio three-dimensional mesh would require many elements to resolve thin plies over a comparatively large in-plane domain. The DG formulation used here instead permits high-aspect-ratio hexahedral elements aligned with the laminate geometry, so that local refinement can be concentrated where it is mechanically useful without forcing the whole domain to be uniformly refined. The earlier validation study showed optimal convergence behaviour on stretched elements, accurate traction transfer across isotropic and orthotropic interfaces, and agreement with established free-edge interlaminar-stress benchmarks while using substantially fewer degrees of freedom than comparable refined three-dimensional models [37].

Within the optimisation workflow, the evaluator maps each candidate laminate design to the interlaminar stress field used for objective construction. The integer-coded design vector is first decoded into the corresponding material stiffnesses and fibre-angle choices; the DG system is then solved under the benchmark-specific prescribed-displacement boundary conditions; and the displacement solution is finally post-processed to recover the relevant stress components. This sequence can be written schematically as

$$\boldsymbol{x} \xrightarrow{\text{decode}} \left\{ \mathbf{D}^{(k)}(\theta_k) \right\}_{k=1}^{n_\ell} \xrightarrow{\text{DG solve}} \mathbf{u}_h(\boldsymbol{x}) \xrightarrow{\text{post-process}} \boldsymbol{\sigma}_{\mathrm{IL},h}(\boldsymbol{x}). \tag{8}$$

Here $\mathbf{u}_h(\boldsymbol{x})$ is the DG displacement solution obtained under the benchmark-specific prescribed-displacement boundary conditions, and $\boldsymbol{\sigma}_{\mathrm{IL},h}(\boldsymbol{x}) = (\sigma_{zz,h}, \tau_{zx,h}, \tau_{yz,h})(\boldsymbol{x})$ collects the interlaminar stress components passed to the objective construction.

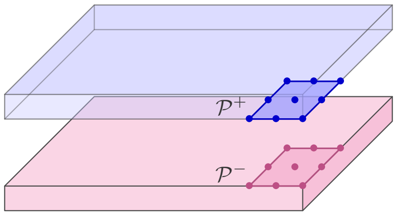


Fig. 2 Schematic of the two-sided interface-patch sampling used for objective construction. The layers are shown in an exploded view only to expose the two interface sides. A square patch and 3×3 sample points are shown for clarity. The sets $\mathcal{P}^-$ and $\mathcal{P}^+$ occupy matching footprints on the two sides of the same interface, with interlaminar stress components extracted at the sampled points

2.3 Objective construction

The optimisation target is constructed from interlaminar stresses sampled over a finite patch, rather than from a single nodal or quadrature-point value. The resulting patch field is then aggregated into a scalar regional measure that provides a consistent stress-severity value for comparing candidate laminates. Figure 2 illustrates the two-sided sampling construction used for this objective.

Let $\mathcal{P}^-$ and $\mathcal{P}^+$ denote the sampled points on the two sides of the monitored interface patch, with $\boldsymbol{r}$ denoting a spatial location in either sampled set. At each sampled point, the relevant interlaminar stress components are the normal stress σ_{zz} and the two interlaminar shear stresses τ_{zx} and τ_{yz}. These components are combined with equal weighting into the local mixed-stress magnitude

$$\phi_s(\boldsymbol{r}) = \left[\sigma_{zz,s}(\boldsymbol{r})^2 + \tau_{zx,s}(\boldsymbol{r})^2 + \tau_{yz,s}(\boldsymbol{r})^2\right]^{1/2}, \qquad s \in \{-,+\}, \tag{9}$$

where s identifies the lower or upper side of the interface patch. The regional response on each side is then defined as the root-mean-square (RMS) value over that side of the patch,

$$R_s = \left(\frac{1}{N_s} \sum_{\boldsymbol{r} \in \mathcal{P}^s} \phi_s(\boldsymbol{r})^2\right)^{1/2}, \qquad s \in \{-,+\}. \tag{10}$$

Here N_s is the number of sampled patch points on side s. Lower values of R_s indicate lower combined interlaminar-stress severity over the monitored region.

This mixed-stress measure is used here as a stress-severity proxy rather than as a material failure criterion. The three interlaminar components are given equal numerical weight so that the objective reflects the magnitude of the local three-dimensional stress state over the monitored patch. It does not account for component-specific strengths, stress-interaction terms, or material-dependent failure envelopes of the kind used in composite failure criteria [52]; these would be needed to convert the present severity measure into a laminate failure index.

Several scalar objective landscapes are then built from these same DG-resolved patch stresses. The purpose is not to change the global optimum, but to test how the

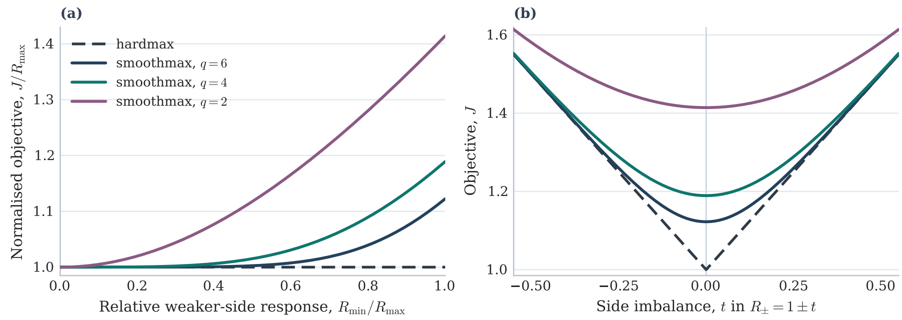


Fig. 3 Illustration of the hardmax and smoothmax side-combination objectives. Panel (a) sets $R_{\max} = 1$ and varies the relative weaker-side response; hardmax remains fixed at $J/R_{\max} = 1$, whereas smoothmax increases as $R_{\min}$ approaches $R_{\max}$. Panel (b) sets $R_{-} = 1 + t$ and $R_{+} = 1 - t$, so the dominant side switches at $t = 0$; hardmax has a sharp kink at this switch, while smoothmax blends the two sides continuously. Smaller q gives a stronger contribution from the non-dominant side

numerical representation of the same interlaminar-stress information affects the optimiser. Two representation choices are separated: how the two sides of the monitored patch are combined, and how the resulting scalar landscape is numerically presented to BO. The hardmax and smoothmax families test conservative worst-side aggregation against smoother side-combination variants, while monotonic transformations provide a controlled way to reshape the objective values while preserving the candidate ranking and global optimum of a fixed base objective. Differences in BO performance can then be interpreted as changes in how easily the optimiser can search the landscape, rather than as changes in the underlying high-fidelity stress data.

The first objective family combines the two patch-side responses using a hard maximum,

$$J_{\mathrm{hard}} = \max(R_{-}, R_{+}). \tag{11}$$

This objective is deliberately conservative: the laminate is judged by the more severe side of the monitored interface. It is therefore a direct worst-side measure of the patch RMS mixed stress. A second family uses a smooth maximum,

$$J_q = \left(R_{-}^{q} + R_{+}^{q}\right)^{1/q}, \qquad q \in \{6, 4, 2\}. \tag{12}$$

Figure 3 illustrates the two numerical effects of replacing the hard maximum with a smooth maximum: the non-dominant patch side contributes to the scalar objective, and switches in the dominant side are rounded rather than kinked.

For larger q, Eq. (12) approaches the hard maximum more closely; for smaller q, it gives a smoother response in which the non-dominant side still contributes more visibly [43]. These smoothmax variants are included because BO performance can depend not only on the physical stress field, but also on the numerical shape of the scalar landscape presented to the surrogate model. By blending the two patch-side responses, the smoothmax can reduce abrupt changes associated with switches in the dominant side and may therefore provide a response surface that is easier for the surrogate to model.

Additional objectives are generated by applying monotonic transformations to the hardmax or smoothmax scalar values. For the hardmax response, a logarithmic transform is included,

$$J_{\log} = \log(J_{\text{hard}} + \epsilon), \qquad \epsilon = 10^{-16}, \tag{13}$$

where the small positive floor avoids evaluating the logarithm at zero. The logarithmic transform is included only for the hardmax baseline as an aggressive-compression reference case, rather than as a full transform family. Power transforms are also applied to both hardmax and smoothmax responses,

$$J_{\alpha,\text{hard}} = J_{\text{hard}}^{\alpha}, \qquad J_{\alpha,q} = J_q^{\alpha}, \qquad \alpha \in \{0.3, 0.5, 0.7\}. \tag{14}$$

These transformations are not new physical definitions of interlaminar stress. They are numerical transformations of the same underlying patch response. Because the logarithmic and power functions used here are monotonic over the positive objective range, they preserve the ordering of candidate designs for a fixed base objective. Their role is therefore not to change the optimum, but to reshape the numerical landscape that BO has to learn. By compressing large objective values and adjusting the contrast among low-objective candidates, these transformations can change the surrogate fit, the resulting acquisition values and, ultimately, how efficiently BO explores the design space [12]. All objective variants used in the benchmark study are treated as minimisation objectives and are summarised in Table 2.

Table 2 Scalar objective variants used in the optimisation benchmarks.

Objective family	Definition	Role
Hardmax	$J_{\text{hard}} = \max(R_-, R_+)$	Conservative worst-side objective.
Smoothmax	$J_q = (R_-^q + R_+^q)^{1/q}$, $q \in \{6, 4, 2\}$	Smooth side-combination variants.
Log transform	$J_{\log} = \log(J_{\text{hard}} + 10^{-16})$	Rank-preserving logarithmic rescaling of the hardmax landscape.
Power transforms	$J_{\alpha,\text{hard}} = J_{\text{hard}}^{\alpha}$ and $J_{\alpha,q} = J_q^{\alpha}$, $\alpha \in \{0.3, 0.5, 0.7\}$	Rank-preserving power rescaling of hardmax and smoothmax landscapes.

2.4 BO strategy

After a scalar patch objective has been defined, the remaining task is to choose which candidate laminates should be evaluated under a limited budget. Each candidate is selected from a finite library of material and fibre-angle choices, and its objective value is obtained only after a DG stress analysis, without access to analytical gradients with respect to the design variables. The resulting problem is therefore a discrete, expensive black-box optimisation problem, making it suitable for BO, which uses a probabilistic surrogate and an acquisition function to allocate evaluations sequentially [16].

The physical objective $J(\boldsymbol{x})$ is a stress-severity quantity to be minimised. The BO implementation uses the equivalent maximisation target

$$y_{\mathrm{BO}}(\boldsymbol{x}) = -J(\boldsymbol{x}), \tag{15}$$

where all reported objective values and success criteria remain in the original minimisation form. After n evaluations, the observed BO data are

$$\mathcal{D}_n = \{(\boldsymbol{x}_i, y_i)\}_{i=1}^{n}, \qquad y_i = y_{\mathrm{BO}}(\boldsymbol{x}_i). \tag{16}$$

Each BO run begins with an initial set of feasible laminate designs, which are evaluated before the first surrogate model is fitted. The sequential stage then updates the surrogate using $\mathcal{D}_n$, selects the next candidate by maximising an acquisition function, evaluates that candidate, and augments the data set.

The surrogate model is a Gaussian process (GP). In the standard scalar-output GP setting, conditioning on the evaluated data gives a normal predictive distribution at an unevaluated design [16, 46],

$$Y(\boldsymbol{x}) \mid \mathcal{D}_n \sim \mathcal{N}\left(\mu_n(\boldsymbol{x}), \sigma_n^2(\boldsymbol{x})\right), \tag{17}$$

where $\mu_n(\boldsymbol{x})$ is the posterior mean and $\sigma_n(\boldsymbol{x})$ is the posterior standard deviation. The posterior mean identifies designs predicted to be favourable, while the posterior uncertainty identifies designs whose response is still poorly resolved. This mean–uncertainty structure is what allows BO to balance exploitation and exploration.

For this finite-library setting, the encoding of the design variables determines how the surrogate model represents similarity between candidate designs and how acquisition values are evaluated over admissible choices [18]. The implementation uses BoTorch [5]: the integer-coded design vectors are passed to a `MixedSingleTaskGP` surrogate with all design-variable columns identified as categorical inputs. The scalar BO targets are standardised during model fitting, and candidate selection is performed over the feasible discrete design set.

The selection criterion is based on expected improvement (EI). For a maximisation target, let the current best observed BO value be

$$y_n^{+} = \max_{1 \leq i \leq n} y_i. \tag{18}$$

The improvement obtained by evaluating a candidate $\boldsymbol{x}$ is then

$$I_n(\boldsymbol{x}) = \left[Y(\boldsymbol{x}) - y_n^{+}\right]_{+}, \tag{19}$$

where $[a]_+$ denotes the positive part of a scalar a, defined as $\max(a, 0)$. The EI is

$$\mathrm{EI}_n(\boldsymbol{x}) = \mathbb{E}\left[I_n(\boldsymbol{x}) \mid \mathcal{D}_n\right]. \tag{20}$$

Here, $\mathbb{E}[\cdot \mid \mathcal{D}_n]$ denotes expectation under the GP posterior predictive distribution after conditioning on the observed data $\mathcal{D}_n$. Equations (18)–(20) express the usual improvement-based BO logic: a candidate is valuable if it has a high predicted value, high uncertainty, or both [16]. In this study, the implemented acquisition is `qLogExpectedImprovement` (`qLogEI`), the logarithmic expected-improvement variant used by BoTorch. The LogEI family reformulates improvement-based acquisition functions to reduce numerical pathologies in standard EI and batch EI, particularly vanishing acquisition values and gradients during acquisition optimisation [3]. The acquisition batch size is $q = 1$, so one laminate is proposed at each BO iteration.

Because the feasible laminate library is finite, the next design is selected by maximising the acquisition function over the feasible unevaluated set,

$$\boldsymbol{x}_{n+1} = \arg \max_{\boldsymbol{x} \in \mathcal{X}_{\text{feas}} \setminus \mathcal{X}_n} \alpha_n(\boldsymbol{x}), \qquad \mathcal{X}_n = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}, \tag{21}$$

where α_n denotes the `qLogEI` acquisition value computed from the current GP posterior. The feasible design set is small enough to be enumerated directly, and already evaluated designs are excluded from the candidate pool. Each BO proposal is therefore a valid laminate from the same finite library used for objective evaluation, without requiring a continuous relaxation or a subsequent projection step. The corresponding BO settings are summarised in Table 3.

Table 3 BO settings used in the benchmark runs.

Item	Setting
Design variables	Integer-coded material and fibre-angle choices from the feasible laminate design space.
Initial design	Random feasible laminates, with sample size set by the benchmark budget.
Surrogate model	BoTorch `MixedSingleTaskGP` with standardised scalar output.
Optimisation target	Maximise $y_{\text{BO}} = -J$ for a physical objective J defined as a minimisation problem.
Acquisition function	`qLogExpectedImprovement` with $q = 1$ and 512 Monte-Carlo samples.
Candidate selection	Enumeration over feasible unevaluated laminates; duplicates are excluded.
Budget convention	Total evaluations equal initial evaluations plus sequential BO evaluations.
Repeated runs	BO is repeated over 10 independent random seeds per objective and budget.

2.5 GA baseline

A GA is used as a population-based derivative-free baseline for the same discrete laminate design space [6]. This choice provides a widely used and transparent reference for discrete stacking-sequence search. The GA operates directly on the same categorical material-angle encoding as BO, requires no gradients or continuous relaxation, and represents the class of evolutionary searches that improve designs through repeated population-level objective queries. It also uses the same feasibility rule and objective definitions as BO, so both optimisers query the same set of admissible laminate designs. Let

$$\mathcal{P}_g = \{\boldsymbol{x}_{g,1}, \dots, \boldsymbol{x}_{g,N_p}\} \subset \mathcal{X}_{\text{feas}} \tag{22}$$

denote the population of N_p feasible designs at generation g. The initial population $\mathcal{P}_0$ is sampled from $\mathcal{X}_{\text{feas}}$. Parent selection uses tournament selection: a subset $\mathcal{T} \subset \mathcal{P}_g$ with $|\mathcal{T}| = \tau$ is sampled, and the selected parent is the best design in that subset,

$$\boldsymbol{x}_{\text{par}} = \arg\min_{\boldsymbol{x} \in \mathcal{T}} J(\boldsymbol{x}). \tag{23}$$

Given two parents $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$, uniform crossover forms an offspring $\boldsymbol{c}$ gene by gene,

$$c_j = \begin{cases} x_j^{(1)}, & z_j = 0, \\ x_j^{(2)}, & z_j = 1, \end{cases} \qquad z_j \sim \text{Bernoulli}(p_{\text{swap}}), \quad j = 1, \dots, d. \tag{24}$$

Mutation is then applied independently to each gene with probability p_{mut}, replacing the selected material or angle code by another valid code for that design variable. The next generation is formed by combining the retained elite set with the newly generated offspring, so that the best designs already found are preserved during population replacement. These implementation choices are standard degrees of freedom in practical GA design, and GA performance can depend on representation, operators, and parameter settings [30].

The main GA settings are summarised in Table 4. The GA evaluates only unique feasible candidates. If crossover or mutation proposes an infeasible or already evaluated design, the candidate is repaired or replaced by another unseen feasible design rather than being assigned a penalised objective value and logged as an evaluation. This convention makes the evaluation budget directly comparable with BO: each logged evaluation corresponds to one distinct feasible laminate objective query.

Table 4 GA settings used for the baseline studies.

Item	Setting
Representation	Integer-coded laminate design vector from Eq. (1).
Initial population	Random feasible designs sampled from $\mathcal{X}_{\text{feas}}$.
Population size	Fixed at 20 individuals per generation.
Selection	Tournament selection with tournament size 3.
Crossover	Uniform crossover with probability 0.9 and per-gene swap probability 0.5.
Mutation	Per-gene mutation probability $1/d$, with $d = 4$ design variables.
Elitism	Best 2 individuals retained between generations.
Candidate filtering	Duplicate or infeasible candidates are repaired or replaced before evaluation.
Budget scan	GA budgets are scanned in 20-evaluation generation increments.
Independent runs	GA is repeated over 10 independent random seeds per objective and budget.

2.6 Offline objective landscapes and performance metrics

The optimisation runs are assessed against the best designs in the same finite laminate library from which they sample candidates. This requires a reference value for every feasible design; although the DG formulation makes the resolved three-dimensional analyses tractable, rerunning the same high-fidelity analysis across repeated BO and GA seeds, budgets and objective variants would be unnecessary duplication. The DG solve is therefore separated from the cheaper objective post-processing. Let b denote a benchmark problem. For each b, the DG displacement field is computed once for every feasible laminate design and stored as

$$\mathcal{U}^{(b)} = \left\{ \left(\boldsymbol{x}, \mathbf{u}_h^{(b)}(\boldsymbol{x}) \right) : \boldsymbol{x} \in \mathcal{X}_{\text{feas}} \right\}, \tag{25}$$

where $\mathbf{u}_h^{(b)}(\boldsymbol{x})$ denotes the discrete DG displacement solution for design $\boldsymbol{x}$ in benchmark b. Different scalar objectives are then obtained by post-processing the stored displacement solutions, rather than by rerunning the DG solver. For objective variant $J_\nu^{(b)}$,

$$J_\nu^{(b)}(\boldsymbol{x}) = \mathcal{Q}_\nu^{(b)} \left(\mathbf{u}_h^{(b)}(\boldsymbol{x}) \right), \tag{26}$$

where $\mathcal{Q}_\nu^{(b)}$ denotes the benchmark-specific extraction of local interlaminar stresses over the monitored patch followed by the objective construction defined in Section 2.3.

The complete reference landscape for objective $J_\nu^{(b)}$ is therefore

$$\mathcal{L}_{J_\nu}^{(b)} = \left\{ \left(\boldsymbol{x}, J_\nu^{(b)}(\boldsymbol{x}) \right) : \boldsymbol{x} \in \mathcal{X}_{\text{feas}} \right\}. \tag{27}$$

This construction gives the global optimum and full ranking of all feasible designs for each benchmark–objective pair. It also makes the repeated BO and GA benchmarking runs inexpensive, since each queried design can be assigned its already computed objective value from $\mathcal{L}_{J_\nu}^{(b)}$. During a BO or GA run, however, the optimiser receives only the objective value of the design it has queried; it is not given access to the full landscape or to the ranking information. The offline landscape is used externally only for objective lookup and for post hoc performance assessment. In the definitions below, J denotes any one scalar objective for a fixed benchmark.

The global optimum for objective J is represented by the set of globally optimal designs

$$\mathcal{X}_J^\star = \arg \min_{\boldsymbol{x} \in \mathcal{X}_{\text{feas}}} J(\boldsymbol{x}), \tag{28}$$

which allows for the possibility that more than one design attains the same minimum objective value. The landscape rank is defined by sorting feasible designs in ascending objective value. The corresponding top-k reference set is denoted by

$$\mathcal{X}_J^{(k)} = \{ \boldsymbol{x} \in \mathcal{X}_{\text{feas}} : \text{rank}_J(\boldsymbol{x}) \leq k \}. \tag{29}$$

For optimiser $m \in \{\text{BO}, \text{GA}\}$, run r, and total evaluation budget B, let

$$\mathcal{H}_{m,r}(B) = \{ \boldsymbol{x}_{m,r,1}, \ldots, \boldsymbol{x}_{m,r,B} \} \tag{30}$$

denote the set of distinct feasible designs evaluated by that run up to budget B. A global-optimum hit is recorded using

$$S^{\star}_{m,r}(B) = \mathbf{1}\left[\mathcal{H}_{m,r}(B) \cap \mathcal{X}^{\star}_{J} \neq \varnothing\right], \tag{31}$$

where $\mathbf{1}[\cdot]$ is the indicator function. Thus, $S^{\star}_{m,r}(B) = 1$ if the run has evaluated at least one globally optimal design by budget B, and $S^{\star}_{m,r}(B) = 0$ otherwise. The same logic is used for top-k performance,

$$S^{(k)}_{m,r}(B) = \mathbf{1}\left[\mathcal{H}_{m,r}(B) \cap \mathcal{X}^{(k)}_{J} \neq \varnothing\right]. \tag{32}$$

The first-hit evaluation is defined as

$$T^{\star}_{m,r} = \min\left\{t : \boldsymbol{x}_{m,r,t} \in \mathcal{X}^{\star}_{J}\right\}, \tag{33}$$

when a global-optimum hit occurs; otherwise it is left undefined for that run. The empirical reliability at budget B is the fraction of independent runs that have reached a globally optimal design,

$$\widehat{\rho}_{m,J}(B) = \frac{1}{R}\sum_{r=1}^{R} S^{\star}_{m,r}(B), \tag{34}$$

where R is the number of independent runs. Because $R = 10$ in the benchmark studies, the reported success rates are empirical fractions with 0.1 resolution and are used for controlled within-study comparisons rather than as population-level reliability estimates.

Equal-budget comparisons report $\widehat{\rho}_{\mathrm{BO},J}(B)$ and $\widehat{\rho}_{\mathrm{GA},J}(B)$ at the same total evaluation budget. The matched-reliability comparison asks a complementary question: for a fixed BO budget B_{BO}, what GA budget is needed to reach at least the same observed reliability? This matched budget is defined as

$$B^{\mathrm{match}}_{\mathrm{GA},J}(B_{\mathrm{BO}}) = \min\left\{B \in \mathcal{B}_{\mathrm{GA}} : \widehat{\rho}_{\mathrm{GA},J}(B) \geq \widehat{\rho}_{\mathrm{BO},J}(B_{\mathrm{BO}})\right\}, \tag{35}$$

where $\mathcal{B}_{\mathrm{GA}}$ is the scanned set of GA budgets. If no scanned GA budget reaches the BO reliability, the case is reported as unmatched over the tested range. This metric gives an evaluation-cost comparison for the selected objectives and benchmark problems, rather than a statement about optimiser performance outside the tested setting.

3 Results and discussion

Both benchmark studies use the same finite material-angle library. The material library is based on the carbon-fibre composite lamina systems compiled by Burhan et al. [7] for the analysis of interlaminar stresses and energy release rates near free edges. Twelve unidirectional lamina systems are retained, with the orthotropic engineering constants listed in Table 5, and the admissible fibre-angle set is $\Theta =$

$\{-70°, -45°, -20°, 0°, 20°, 45°, 60°, 75°, 90°\}$. This finite angle library is used to define a controlled benchmark search space: it contains the conventional 0°, ±45° and 90° directions together with additional off-axis angles, and its deliberately non-uniform, mildly asymmetric spacing avoids reducing the search to a perfectly mirrored angle grid.

Table 5 Material-property library used for the candidate lamina choices. Elastic and shear moduli are in GPa; Poisson ratios are dimensionless. Material data are compiled from Burhan et al. [7].

Material	E_x	$E_y = E_z$	$G_{xy} = G_{zx}$	G_{yz}	$\nu_{xy} = \nu_{xz}$	ν_{yz}
IM6/3501-6	145.00	9.65	5.20	3.33	0.30	0.45
IM7/977-3	164.00	8.97	5.00	3.01	0.32	0.49
T300/5208	181.00	10.30	7.17	3.63	0.27	0.42
IM7/MTM45-1	145.00	10.30	5.30	3.93	0.30	0.31
T800/924C	168.00	9.50	4.60	3.74	0.27	0.27
T800/3900	142.00	7.79	4.00	2.55	0.34	0.53
T650-35/PMR-15	127.00	8.25	4.20	2.81	0.41	0.47
T700/M21	131.00	8.20	4.30	3.04	0.35	0.35
T300/970	135.00	8.00	5.00	2.99	0.25	0.34
T700/MTM57	128.00	7.90	3.00	2.82	0.30	0.40
HTS/977-2	139.00	8.80	4.70	2.97	0.26	0.48
IM7/5250-4	166.00	10.34	5.79	3.31	0.31	0.56

With this shared library, the benchmark studies instantiate the general design-space definition with $n_\ell = 2$, $n_m = 12$ and $n_\theta = 9$, giving $|\mathcal{X}_{\mathrm{raw}}| = 12^2 \times 9^2 = 11664$ possible integer-coded configurations before feasibility filtering. For this two-group benchmark library, Eq. (3) becomes

$$\mathcal{X}_{\mathrm{feas}} = \{\boldsymbol{x} \in \mathcal{X}_{\mathrm{raw}} : (m_1, \theta_1) \neq (m_2, \theta_2)\}\,. \tag{36}$$

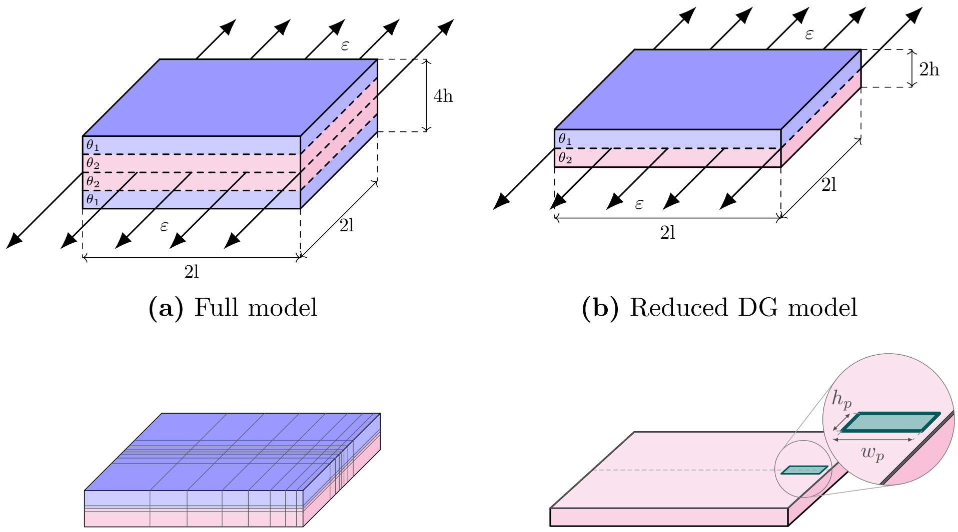
This excludes only identical material-angle group pairs while retaining designs with the same material but different fibre angles, and designs with different materials but the same fibre angle. No additional manufacturing-compatibility filter is imposed between different material systems. In practice, some mixed-system laminates could require further restrictions because the entries in Table 5 represent complete lamina systems with their own matrix and processing route, and not all resin systems would necessarily be co-curable in one physical laminate. Those restrictions are deliberately left outside the present benchmark so that the study isolates optimiser behaviour and local-stress response over a broad elastic-property library. They are nevertheless straightforward to include in the same framework: incompatible material pairs can be removed by adding a material-pair compatibility mask to the feasibility rule in Eq. (3), as in manufacturing-constrained laminate optimisation workflows [17]. The resulting benchmark design space contains $|\mathcal{X}_{\mathrm{feas}}| = 11556$ candidate laminates.

3.1 Free-edge benchmark

The first benchmark considers the straight free-edge laminate problem as a canonical stack-driven interlaminar-stress case for testing the DG-BO framework. The benchmark geometry, prescribed-extension loading, reduced DG domain, and graded high-aspect-ratio hexahedral mesh follow the free-edge validation configuration of the earlier DG study [37]; the present work keeps these analysis settings fixed while replacing the fixed layup with material and fibre-angle choices sampled from the design library. The shared feasible laminate library is defined at the start of this section, while the objective definitions, offline landscape construction, and performance metrics are those defined in Section 2. Because the same finite design library is used throughout, differences in success rate, final rank, and first-hit behaviour can be interpreted as effects of the objective landscape and optimisation algorithm (BO or GA), rather than as effects of changing the available candidate designs. Although all 17 free-edge objective variants defined in Section 2.3 were evaluated, this section retains five representative cases spanning hardmax, smoothmax, logarithmic scaling and square-root transformations; the full BO and GA results for all objectives and budgets are reported in Appendix A, Table 12.

Figure 4 summarises the physical and numerical setup used for this benchmark. The full specimen is a symmetric laminate with in-plane dimensions $2l \times 2l$ and total thickness $4h$, with $l = 8h$. The loading is imposed through prescribed displacement boundary conditions corresponding to a uniform macroscopic tensile strain $\varepsilon_0 = 1 \times 10^{-6}$ indicated by ε in Fig. 4(a). Under this prescribed extension, local out-of-plane stresses develop near the free boundary because adjacent laminate groups can have different orthotropic stiffnesses and fibre orientations [32, 44]. Symmetry reduces the DG analysis to the upper half model in Fig. 4(b), which contains the two laminate groups associated with the two material and fibre-angle pairs used in the design vector for the numerical studies. The corresponding hexahedral DG mesh is shown in Fig. 4(c); refinement is concentrated towards the free edge and the specimen midline, placing the finest elements around the monitored interface patch from which the interlaminar stresses are sampled, while stretched elements with a maximum aspect ratio of approximately 40 keep the mesh coarse away from this region. The monitored patch is shown separately in Fig. 4(d), with the upper layer hidden only to expose the interface plane. The patch lies next to the free edge with offset $\epsilon_p = 0.1h$, width $w_p = 3h$, and transverse height $h_p = 0.1l$, and is centred along the specimen midline away from the corners. The offset keeps the sampled region within the near-edge boundary layer without defining the objective directly at the free-edge boundary, where the local singular response is most pronounced [45]. The midline placement limits corner effects [31], so the scalar objectives primarily reflect the stack-driven free-edge response produced by material choice, fibre angle, and stacking arrangement.

Table 6 identifies the landscape targets against which the optimiser runs are judged. For all five retained objectives, the same three laminate configurations occupy the top-three positions. The global optimum uses IM7/MTM45-1 at 90° in group 1 and IM6/3501-6 at 90° in group 2; the second-ranked design swaps the two material assignments while keeping both angles at 90°; and the third-ranked design combines T700/MTM57 at 90° in group 1 with IM7/MTM45-1 at 90° in group 2. The reversed order of the first two material pairings gives distinct, but very close, objective values, showing that the group assignment is not completely interchangeable in the free-edge response. The common top-three set is expected within each monotonic transform



(a) Full model **(b)** Reduced DG model

(c) DG mesh **(d)** Monitored patch

Fig. 4 Free-edge benchmark setup: (a) full symmetric laminate model, (b) reduced model used for DG evaluation, (c) free-edge DG mesh containing 256 hexahedral elements, and (d) monitored interface patch used for objective construction, with the inset defining the local patch dimensions w_p and h_p. The patch panel hides the upper layer only for visualisation

Table 6 Top-three free-edge landscape designs for the retained objective variants defined in Table 2.

Objective	Rank	Group 1 material	θ_1	Group 2 material	θ_2	Objective value
hardmax	1	IM7/MTM45-1	90°	IM6/3501-6	90°	8.036×10^{-14}
	2	IM6/3501-6	90°	IM7/MTM45-1	90°	8.136×10^{-14}
	3	T700/MTM57	90°	IM7/MTM45-1	90°	8.743×10^{-14}
log(hardmax)	1	IM7/MTM45-1	90°	IM6/3501-6	90°	-30.1510
	2	IM6/3501-6	90°	IM7/MTM45-1	90°	-30.1387
	3	T700/MTM57	90°	IM7/MTM45-1	90°	-30.0668
pow0.5(hardmax)	1	IM7/MTM45-1	90°	IM6/3501-6	90°	2.835×10^{-7}
	2	IM6/3501-6	90°	IM7/MTM45-1	90°	2.852×10^{-7}
	3	T700/MTM57	90°	IM7/MTM45-1	90°	2.957×10^{-7}
smoothmax($q = 4$)	1	IM7/MTM45-1	90°	IM6/3501-6	90°	8.970×10^{-14}
	2	IM6/3501-6	90°	IM7/MTM45-1	90°	9.104×10^{-14}
	3	T700/MTM57	90°	IM7/MTM45-1	90°	1.022×10^{-13}
pow0.5(smoothmax($q = 4$))	1	IM7/MTM45-1	90°	IM6/3501-6	90°	2.995×10^{-7}
	2	IM6/3501-6	90°	IM7/MTM45-1	90°	3.017×10^{-7}
	3	T700/MTM57	90°	IM7/MTM45-1	90°	3.196×10^{-7}

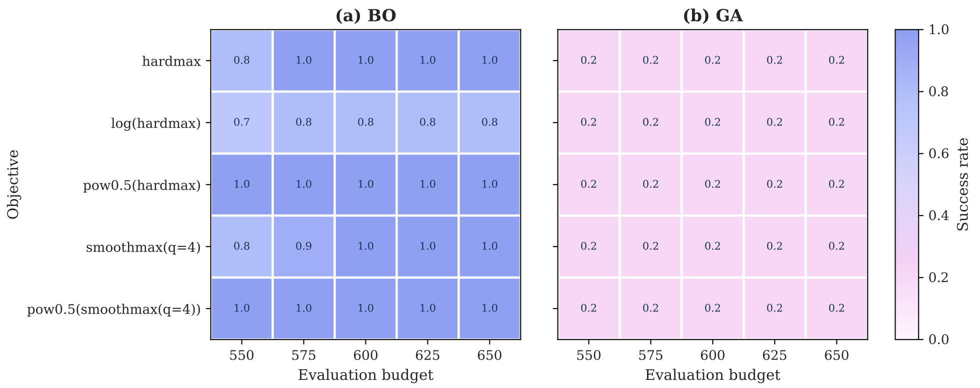


Fig. 5 Empirical global-optimum success rates for BO and GA in the free-edge benchmark. Each cell reports the fraction of 10 independent runs that found the exact global optimum for the corresponding objective and evaluation budget; success-rate entries therefore have 0.1 resolution

family, since the logarithmic and power transforms preserve the ranking of a given positive base objective. By contrast, hardmax and smoothmax are different objective definitions, so any overlap between their top-ranked designs is an observed result rather than an enforced property. The table therefore separates the objective variants even when the ranked designs coincide: the optimiser sees different numerical objective values, and the subsequent reliability comparison tests how those different scalar landscapes affect search performance.

The all-90° top-three set should not be interpreted as an angle-only rule. A direct check within the hardmax landscape shows that the preferred angle can change strongly with the material pairing. For example, the ordered pair T700/MTM57–T800/3900 is ranked 11th when both groups use 0° ($J = 1.591 \times 10^{-9}$), but only 329th when the same ordered pair uses 90° ($J = 6.906 \times 10^{-8}$). Similarly, HTS/977-2–IM6/3501-6 ranks 13th at 45°/45° ($J = 5.847 \times 10^{-9}$), whereas the corresponding 90°/90° design ranks 368th ($J = 7.765 \times 10^{-8}$). These comparisons show that the free-edge optimum is not a generic consequence of choosing 90° plies; it depends on how fibre angle and material compatibility interact in the local stress response [32]. A broader post-hoc landscape analysis, reported in Appendix D, supports the same interpretation across the full hardmax landscape.

With these reference designs and their landscape context established, the first optimisation question is the empirical reliability of BO under the strict global-optimum criterion. Here, reliability means the global-optimum success rate; the fraction of ten independent runs whose best evaluated design is the exact optimum at a given evaluation budget. In the free-edge BO runs, the 550–650 total-evaluation budgets use a fixed 500 acquisition steps and initial random designs of 50, 75, 100, 125 and 150, respectively. Figure 5 reports this quantity for the five objective variants retained for the main-text comparison.

The BO panel shows that this exact-hit success rate is strongly objective dependent, even though the objectives are constructed from the same DG stress fields over the same monitored patch. The two 0.5-power objectives, applied to hardmax and smoothmax($q = 4$), reach a success rate of 1.0 for every budget from 550 to 650 evaluations. These transformed objectives are therefore not only monotonic transformations of the underlying patch-stress measures, but also the most robust BO targets among the selected free-edge cases. The raw objectives also perform well, but with a clearer dependence on budget. The raw hardmax objective reaches a success rate of 0.8 at 550 evaluations and 1.0 at all larger tested budgets, while the raw smoothmax($q = 4$) objective increases from 0.8 at 550 evaluations to 0.9 at 575 evaluations and 1.0 from 600 evaluations onward. The weakest selected BO case is the logarithmic hardmax transform, which remains between 0.7 and 0.8 success over the same budget range.

The GA panel provides the corresponding baseline under the same objective definitions and evaluation budgets. In contrast to BO, the GA success rate remains at 0.2 for every selected objective and budget. The identical GA results for monotonic transforms of the same base objective (for example, hardmax and pow0.5(hardmax)) are expected because the GA implementation uses tournament selection and population sorting based on objective ordering; applying a strictly monotonic transform to a positive base objective preserves that ordering. The heatmap also shows that changing the base objective from hardmax to smoothmax($q = 4$) does not improve GA exact-hit reliability across the selected budgets. The objective dependence in Fig. 5 is therefore mainly a BO effect: the surrogate and acquisition strategy respond differently to different scalar representations of the same DG-derived stress response.

The success-rate results also separate the effect of changing how the two sides of the monitored patch are combined from the effect of numerically rescaling the resulting objective. Replacing hardmax by smoothmax($q = 4$) changes the scalar landscape, but is not by itself the decisive improvement in this free-edge benchmark: the raw hardmax and raw smoothmax objectives show broadly similar BO reliability trends, whereas both 0.5-power objectives reach complete success across the tested budgets. The main gain therefore comes from the power-rescaled numerical presentation of the objective, with smoothmax acting as a compatible smooth aggregation rather than the dominant source of the performance gain.

The comparison between the logarithmic and 0.5-power transforms is especially informative because both are monotone transformations of positive base objectives. They preserve the identity of the global optimum, but change the numerical spacing of the objective values that the GP surrogate must model. The 0.5-power map is concave, so it compresses the high-objective tail associated with clearly poor designs while giving greater numerical resolution to differences among low-objective candidates. Although the BO targets are standardised before GP fitting, this nonlinear redistribution remains: the surrogate is fitted to a response that is less dominated by large stress values, and the acquisition function can more clearly distinguish small improvements near the optimum. By contrast, the logarithmic transform appears too aggressive for this landscape, preserving the ranking but over-compressing or distorting the numerical region that separates near-optimal candidates. Thus, the improved reliability of the 0.5-power objectives is not caused by redefining the best laminate; it is best interpreted as a surrogate-modelling and acquisition effect that makes the same optimum easier for BO to identify.

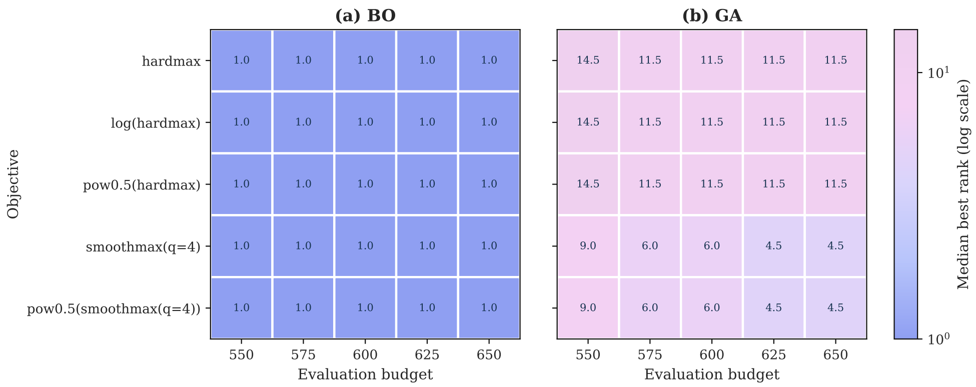


Fig. 6 Median final best rank for BO and GA in the free-edge benchmark. Each cell reports the median rank of the best final design over 10 independent runs for the corresponding objective and evaluation budget. Rank 1 denotes the global optimum of the corresponding objective landscape; larger ranks indicate lower positions in the landscape ordering

Exact global-optimum success is a deliberately strict binary criterion: a run either finds the best design in the finite objective landscape or it does not. Figure 6 reports the final best rank, which gives useful context for that binary criterion by showing where the best design found by each run lies in the full objective ordering.

The BO median final rank is 1.0 for every selected objective and budget, including the log(hardmax) and low-budget raw-objective cases for which the success rates in Fig. 5 are below one. Thus, the reduced BO success rates for those cases are caused by a minority of missed runs rather than by a shift in the typical BO outcome away from the global optimum. The GA panel shows a different pattern. Because the monotonic transforms preserve the GA fitness ordering, hardmax, log(hardmax), and pow0.5(hardmax) give identical GA median ranks: 14.5 at 550 evaluations and 11.5 for all larger tested budgets. The smoothmax($q = 4$) and pow0.5(smoothmax($q = 4$)) cases also match each other, improving with budget from median rank 9.0 at 550 evaluations to 4.5 at 625 and 650 evaluations. Even in these smoother cases, however, the typical GA run does not reach the global optimum and remains outside the top-three set. The rank evidence therefore supports the strict success-rate result while adding context: BO is typically ending at the best design, whereas GA more often ends in a lower-ranked region of the same objective landscape.

Figure 7 closes this strict reliability comparison by putting the success-rate and rank evidence on the same paired scale. Each objective-specific marker represents one objective at one budget. The left panel reports the paired success-rate difference, $\mathrm{SR}_{\mathrm{BO}} - \mathrm{SR}_{\mathrm{GA}}$, while the right panel reports the paired median-rank difference, $\widetilde{r}_{\mathrm{GA}} - \widetilde{r}_{\mathrm{BO}}$, so positive values indicate a BO advantage in both panels.

All paired points lie on the BO-favouring side of both panels, and the mean markers remain well separated from zero at every budget. The figure therefore summarises the main equal-budget result of the free-edge study: under the same feasible design library, objective landscapes, and number of DG-derived objective queries, BO is both more likely to recover the exact global optimum and more likely to finish at a higher-ranked laminate than GA. This establishes BO as the more reliable search layer for the strict

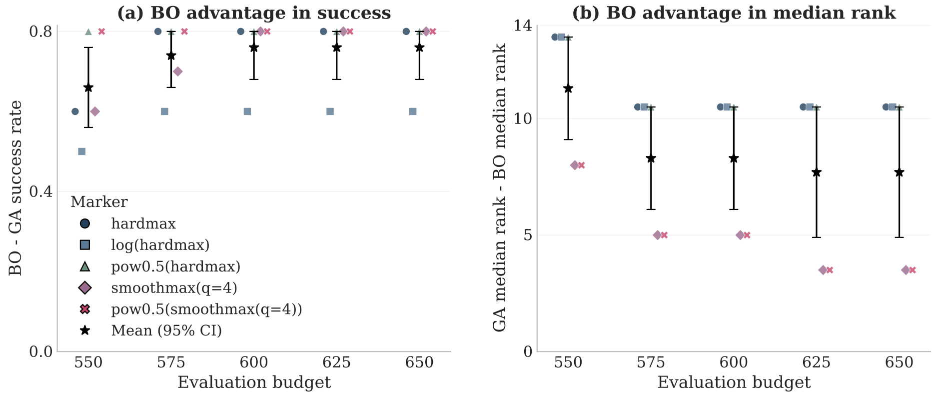


Fig. 7 Equal-budget BO–GA comparison for the free-edge benchmark. Objective-specific markers show paired differences at each evaluation budget, with small horizontal offsets used only to avoid overlap; central stars and error bars show the mean across the selected objectives and its 95% bootstrap confidence interval. Positive values indicate higher BO success rate in the left panel and better BO median final rank in the right panel

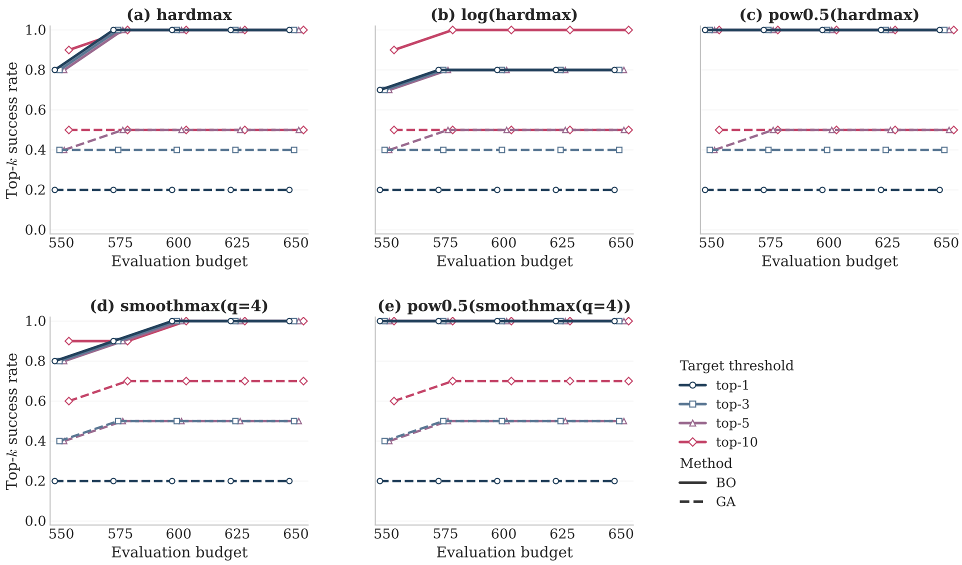


Fig. 8 Top-k target success rates for the free-edge benchmark. A run is counted as successful for a top-k target if the final best design has rank k or better in the corresponding objective landscape. Solid lines denote BO and dashed lines denote GA; marker shapes and small horizontal offsets are used only to expose coincident target-threshold curves

rank-1 target in this benchmark. The next step is to relax that target from a single global optimum to small top-ranked sets, so that near-optimal terminal outcomes can be distinguished from weaker misses. Figure 8 reports this target-ladder measure for $k \in \{1, 3, 5, 10\}$.

For the two 0.5-power objectives, BO reaches the top-1 design in every run at every tested budget, so all target thresholds coincide at a success rate of 1.0. For the raw hardmax objective, broadening the target to top-10 increases the 550-evaluation BO success rate from 0.8 to 0.9, and all thresholds reach 1.0 from 575 evaluations onward. The raw smoothmax($q = 4$) objective follows a similar pattern, reaching top-10 success of 0.9 at 550 and 575 evaluations and complete success from 600 evaluations onward.

The log(hardmax) objective shows most clearly why the target threshold matters. Its exact-hit success is only 0.7–0.8, and the top-3 and top-5 rates are identical to the top-1 rate because the missed runs do not end at ranks 2–5. However, the top-10 rate increases to 0.9 at 550 evaluations and 1.0 for all larger tested budgets. Thus, the weaker BO cases are not usually failing to find a competitive region of the landscape; rather, their missed runs tend to land outside the very small top-three and top-five sets. The dashed GA curves provide a reference for the same target definition: GA reaches the exact optimum in only 0.2 of runs for these objectives, and even the top-10 success rate remains between 0.5 and 0.7. This rank-threshold view strengthens the strict success-rate result: BO is not only more likely to hit the single best laminate, but also more consistently reaches the best region of the DG-generated objective landscape.

The preceding results describe where each run ends; the next question is when the global optimum is first found. Figure 9 shows Kaplan–Meier first-hit curves at the representative 600-evaluation budget. For an evaluation count b, the plotted survival estimate is

$$\widehat{S}(b) = \widehat{P}(B_{\text{hit}} > b) = \prod_{u_i \le b} \left(1 - \frac{d_i}{n_i}\right),$$

where B_{hit} is the first-hit evaluation count, u_i are the distinct observed first-hit counts, d_i is the number of runs that first find the global optimum at u_i, and n_i is the number of runs that have not yet found the optimum immediately before u_i. Runs that do not find the optimum within the evaluation budget are treated as censored. The vertical axis therefore estimates the probability that a run has not yet found the global optimum by a given evaluation count; a faster and more reliable optimisation process appears as a curve that drops earlier and approaches zero. This first-hit view separates objectives that ultimately reach the same success rate but do so at different points in the evaluation budget.

The two 0.5-power objectives are again the strongest BO targets. The median first-hit evaluations are 200 for pow0.5(hardmax) and 204 for pow0.5(smoothmax($q = 4$)), with all ten runs successful by evaluations 274 and 323, respectively. The raw hardmax and raw smoothmax($q = 4$) objectives also reach complete BO success at this budget, but later: their median first-hit evaluations are 316.5 and 332, and the latest successful hits occur at evaluations 565 and 414, respectively. The log(hardmax) transform is both slower and less reliable, with eight successful BO runs, a median first hit of 437 among those successful runs, and two censored runs that have not found the optimum by the end of the 600-evaluation budget.

The GA curves reinforce that the difference is not only in final success rate but also in first-hit dynamics. For each selected objective at 600 evaluations, only two of the ten GA runs find the global optimum. The first-hit times among those successful GA runs can occur early, especially for the smoothmax cases, but the curves remain high

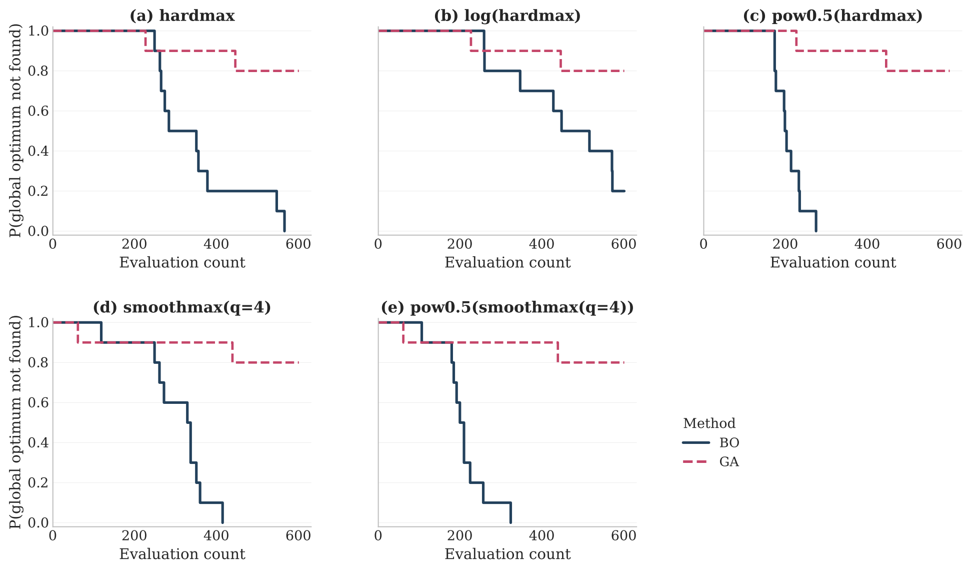


Fig. 9 Kaplan–Meier first-hit curves for the free-edge benchmark at a 600-evaluation budget. The vertical axis gives the estimated probability, over the 10 independent runs, that a run has not yet found the global optimum by the corresponding evaluation count; lower curves therefore indicate earlier and more reliable discovery

because most GA runs are censored without an exact hit. The free-edge benchmark therefore indicates that the transformed BO objectives are not merely more likely to reach the optimum; they also make the global optimum easier for BO to locate early in the sequential search. For a high-fidelity DG evaluator, this earlier first hit is practically important because a useful low-stress laminate is identified after fewer high-fidelity objective queries.

The equal-budget comparison answers whether BO and GA perform differently when both are given the same number of objective evaluations. A complementary question is how many GA evaluations are required to match the reliability already obtained by BO. This is the matched-reliability comparison defined in Eq. (35). For the results of the free-edge benchmark, the BO reference is fixed at 600 total evaluations, and the target reliability for each objective is the empirical BO global-optimum success rate over the ten independent runs. The GA budget is then increased from 600 evaluations in steps of one GA generation, corresponding to 20 additional candidate evaluations, and the first budget reaching at least the BO success rate is recorded.

The matched-reliability results are summarised in Table 7. The four objectives for which BO reaches complete success at 600 evaluations require 3080 GA evaluations for hardmax and pow0.5(hardmax), and 3300 GA evaluations for smoothmax($q = 4$) and pow0.5(smoothmax($q = 4$)). For log(hardmax), the BO target success rate at 600 evaluations is 0.8, and GA reaches the same reliability at 2420 evaluations. The matched-reliability ratios therefore range from 4.03 to 5.50 times the BO budget.

The mean first-hit evaluations in Table 7 are computed over successful runs only: the log(hardmax) values are averaged over eight successful runs for each method, while the other selected objectives are averaged over ten successful runs. Even after increasing the GA budget to match BO reliability, the successful GA runs find the optimum later than the successful BO runs for every selected objective, with first-hit ratios from 3.15 for log(hardmax) to 7.95 for pow0.5(hardmax).

The evaluation-count savings can also be expressed as estimated computational time. Because the optimiser benchmarks use stored objective values rather than rerunning the DG solver at every query, the estimate separates the recorded optimiser/lookup overhead from the assumed cost of a high-fidelity DG objective evaluation. For method $m \in \{\mathrm{BO}, \mathrm{GA}\}$ and objective J, the mean time to first hit is estimated as

$$\widehat{C}_{m,J} = \overline{O}^{\star}_{m,J} + \overline{T}^{\star}_{m,J} c_{\mathrm{DG}},$$

where $\overline{T}^{\star}_{m,J}$ is the mean first-hit evaluation count over successful runs, $\overline{O}^{\star}_{m,J}$ is the corresponding recorded optimiser/lookup overhead, c_{DG} is the cost of one high-fidelity DG objective evaluation, and $\widehat{C}_{m,J}$ is the estimated mean time to first hit. Here c_{DG} is set to 8 s, the average per-design DG evaluation time reported in the earlier DG study [37]. The same model also gives a break-even DG evaluation cost,

$$c^{\mathrm{BE}}_{\mathrm{DG},J} = \frac{\overline{O}^{\star}_{\mathrm{BO},J} - \overline{O}^{\star}_{\mathrm{GA},J}}{\overline{T}^{\star}_{\mathrm{GA},J} - \overline{T}^{\star}_{\mathrm{BO},J}},$$

at which the overhead difference exactly offsets the difference in mean first-hit evaluation count, giving BO and the matched GA the same estimated mean time to first hit for objective J. Under the 8 s DG-evaluation assumption, BO reaches the optimum in an estimated mean time of 30.5–69.3 min, whereas the matched-GA estimates range from 178.4 to 219.6 min. The corresponding GA/BO time ratios are 2.57–7.16. This is the practical consequence of using BO as the decision layer around the DG evaluator: the same high-fidelity stress analysis is queried far fewer times before the target reliability is reached. The break-even values in Table 7 are only 0.13–0.84 s per evaluation, well below the DG cost assumed here, so GA would become time competitive only if each objective query were almost a lightweight lookup rather than a high-fidelity DG solve.

Table 7 Matched-reliability summary for the free-edge benchmark. The BO reference budget is 600 evaluations. Ratios are reported as GA/BO. First-hit and time values are means over successful runs only. Estimated time assumes 8 s per high-fidelity objective evaluation plus recorded lookup/optimiser overhead. The break-even cost is the DG evaluation time at which BO and the matched GA have the same mean time to first hit, despite different mean first-hit counts.

Objective	BO target success	Budget			Mean first hit			Time to first hit [min]			Break-even (s/eval)
		BO	GA	GA/BO	BO	GA	GA/BO	BO	GA	GA/BO	
hardmax	1.0	600	3080	5.13	352.2	1646.9	4.68	57.6	219.6	3.82	0.49
log(hardmax)	0.8	600	2420	4.03	424.1	1337.8	3.15	69.3	178.4	2.57	0.84
pow0.5(hardmax)	1.0	600	3080	5.13	207.1	1646.9	7.95	30.7	219.6	7.16	0.13
smoothmax($q = 4$)	1.0	600	3300	5.50	302.0	1408.4	4.66	46.9	187.8	4.00	0.36
pow0.5(smoothmax($q = 4$))	1.0	600	3300	5.50	208.0	1408.4	6.77	30.5	187.8	6.16	0.14

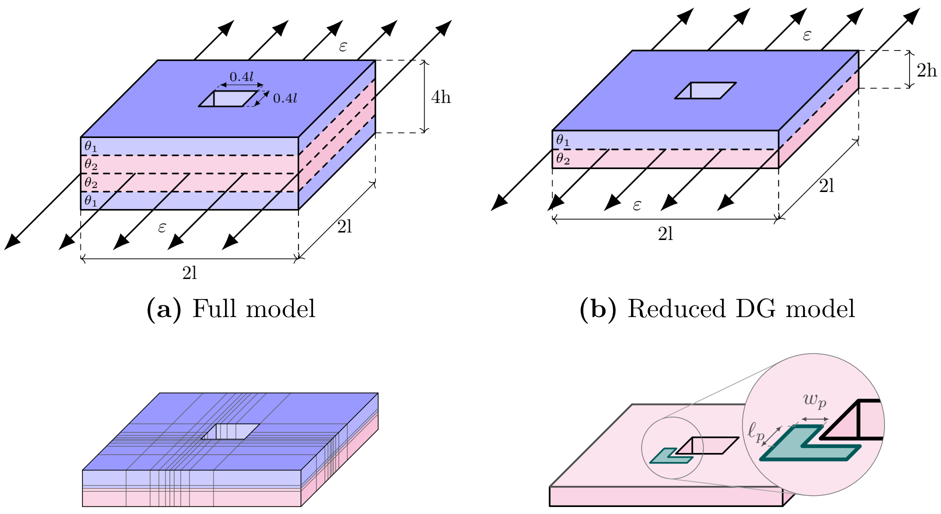
3.2 Cutout benchmark

The second benchmark examines the same DG-BO framework in a laminate containing a centred rectangular cutout. Unlike the straight free-edge case, the monitored interlaminar stresses are now associated with a local geometric discontinuity, so the critical stress field is driven by the interaction between material choice, fibre orientation, stacking arrangement, and cutout-corner geometry [25]. The feasible laminate library, objective-construction procedure, and optimiser comparison metrics are kept the same as in the free-edge benchmark. Differences between the free-edge and cutout benchmark outcomes can therefore be interpreted as the effect of moving from a canonical straight-edge stress mechanism to a more localised cutout-corner stress concentration.

Figure 10 summarises the physical and numerical setup used for the cutout benchmark. The full model in Fig. 10(a) has in-plane dimensions $2l \times 2l$ and total thickness $4h$, with $l = 8h$. As in the free-edge benchmark, the loading is imposed through prescribed displacement boundary conditions corresponding to a macroscopic tensile strain $\varepsilon_0 = 1 \times 10^{-6}$, indicated by ε in the figure. A through-thickness rectangular cutout is placed at the specimen centre, with side length $0.4l$ in both in-plane directions. Symmetry reduces the DG analysis to the upper half model shown in Fig. 10(b), which contains the two laminate groups represented by the two material and fibre-angle pairs in the design vector. The cutout objective landscapes are generated with the one-corner-refined mesh in Fig. 10(c), where the in-plane grid is concentrated around the monitored cutout corner and through-thickness refinement is retained around the laminate interface. As in the free-edge mesh, stretched hexahedral elements with a maximum aspect ratio of approximately 48 allow this local refinement while keeping the mesh coarse away from the corner region. Figure 10(d) shows the monitored L-shaped patch on the exposed interface plane, with the upper layer hidden only for visualisation; the patch is placed around the selected cutout corner with offset $\epsilon_p = 0.1h$, arm width $w_p = 1.5h$, and arm length $\ell_p = 0.2l$. As in the free-edge patch definition, the offset keeps the extraction region inside the local stress-concentration zone while avoiding an objective defined directly at the geometric discontinuity, where the corner response is most singular and mesh-sensitive [25].

Table 8 Top-three designs in the cutout benchmark landscape for the retained hardmax objective and its 0.5-power transform.

Objective	Rank	Group 1 material	θ_1	Group 2 material	θ_2	Objective value
hardmax	1	T800/924C	45°	T700/MTM57	45°	4.815×10^{-7}
	2	T800/924C	20°	T700/MTM57	20°	5.055×10^{-7}
	3	T800/924C	45°	T700/M21	45°	5.093×10^{-7}
pow0.5(hardmax)	1	T800/924C	45°	T700/MTM57	45°	6.939×10^{-4}
	2	T800/924C	20°	T700/MTM57	20°	7.110×10^{-4}
	3	T800/924C	45°	T700/M21	45°	7.137×10^{-4}



(a) Full model **(b)** Reduced DG model

(c) One-corner DG mesh **(d)** Monitored patch

Fig. 10 Cutout benchmark setup: (a) full symmetric laminate model with a centred rectangular cutout, (b) reduced model used for DG evaluation, (c) one-corner refined DG mesh containing 336 hexahedral elements and used to generate the cutout objective landscapes, and (d) monitored L-shaped interface patch used for objective construction, with the inset defining the local patch dimensions. The patch panel hides the upper layer only for visualisation

Table 8 defines the cutout benchmark landscape targets used in the subsequent optimiser comparisons. The cutout study is restricted to two hardmax-based objectives. Raw hardmax is retained as the original local peak-stress objective, while pow0.5(hardmax) is retained because it gave the strongest BO performance in the free-edge benchmark across exact-hit reliability, final-rank evidence, and first-hit dynamics. Comparing the two therefore keeps a direct raw-objective reference while testing whether the most effective free-edge transformation also improves BO performance for the cutout-corner stress concentration. For the retained hardmax objective, the global optimum combines T800/924C at 45° in group 1 with T700/MTM57 at 45° in group 2. The next two designs either retain the same material pair at 20° or replace the second-group material with T700/M21 while keeping both groups at 45°. The same three designs appear in the same order for pow0.5(hardmax), as expected because the 0.5-power transform is strictly monotonic over the positive hardmax objective values. Thus, the transformed objective does not redefine the best laminate for the cutout benchmark; it changes the numerical spacing of the objective values that BO models while preserving the finite-landscape ordering. These leading designs also show that the cutout-corner stress mechanism favours same-angle off-axis families rather than simply reusing the free-edge 90°/90° family; the full post-hoc landscape analysis supporting this interpretation is reported in Appendix D.

The cutout benchmark uses a much tighter range of 100–200 total evaluations to examine BO and GA performance when the available laminate-evaluation budget is strongly limited. For BO, this range includes 50 initial samples followed by the

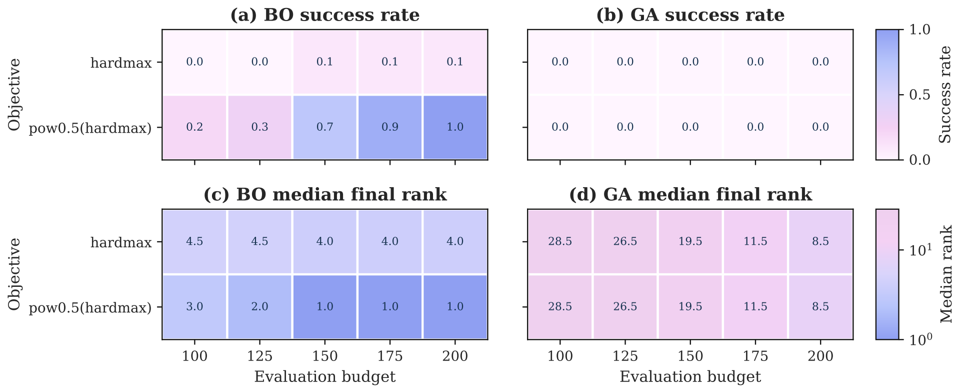


Fig. 11 Cutout benchmark heatmaps for empirical exact-optimum success rate and median final rank over the tight budget range of 100–200 total evaluations. The upper row reports the fraction of 10 independent runs that found the exact optimum for each objective and evaluation budget; success-rate entries therefore have 0.1 resolution. The lower row reports the median rank of the best final design over the same runs. Rank 1 denotes the global optimum; larger ranks indicate lower-ranked designs in the corresponding cutout objective landscape

remaining sequential acquisitions. Figure 11 reports the corresponding exact-optimum success rates and median final ranks. Under this restricted budget, the raw hardmax objective is difficult for BO because the surrogate must learn the untransformed peak-stress ordering from relatively few observations: no run reaches the global optimum at 100 or 125 evaluations, and only one of ten runs does so from 150 evaluations onward. The final-rank evidence gives a more complete view of these misses. The BO median final rank for hardmax is 4.5 at 100–125 evaluations and 4.0 at 150–200 evaluations, so the search usually moves towards the best part of the cutout benchmark landscape but rarely reaches the exact rank-1 design within this budget window.

The 0.5-power transform changes this behaviour without changing the global target identified in Table 8. For pow0.5(hardmax), BO reaches the global optimum in 2 out of 10 runs at 100 evaluations, 3 out of 10 at 125 evaluations, 7 out of 10 at 150 evaluations, 9 out of 10 at 175 evaluations, and all 10 runs at 200 evaluations. Its median final rank also reaches 1.0 from 150 evaluations onward, showing that the rescaled objective values make the cutout benchmark optimum much easier for BO to identify within the tight evaluation budget.

The GA baseline shows that this budget range is demanding for a population-based search. At the same total evaluation budgets, GA does not reach the exact global optimum for either retained objective variant. Its median final rank improves from 28.5 at 100 evaluations to 8.5 at 200 evaluations, but the typical GA run remains outside the top-three design set. The GA entries are identical for hardmax and pow0.5(hardmax), because the power transform preserves the hardmax ordering used for selection. Together, the BO and GA results establish the main cutout benchmark result at this stage: the transformed BO target reaches complete exact-optimum reliability by 200 evaluations, whereas equal-budget GA improves in rank but has not yet reached the global optimum. The next comparison broadens the target from a single global design to top-ranked design sets.

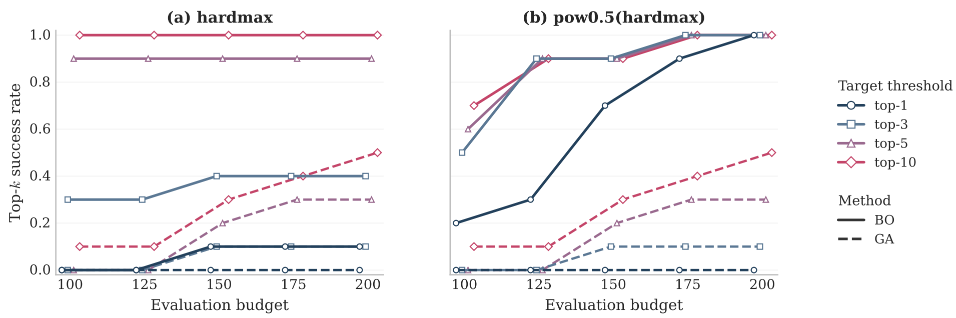


Fig. 12 Top-k target success rates for the cutout benchmark. A run is counted as successful for a top-k target if the final best design has rank k or better in the corresponding objective landscape. Solid lines denote BO and dashed lines denote GA; marker shapes and small horizontal offsets are used only to expose coincident target-threshold curves

Figure 12 applies the top-k success definition in Eq. (32) to the retained cutout objectives, with $k \in \{1, 3, 5, 10\}$. This broader target set separates two different outcomes: a run may miss the exact global optimum but still enter the near-optimal part of the finite landscape, or it may remain in a lower-ranked region.

For the raw hardmax objective, the target ladder shows that BO is already locating competitive designs even when the strict exact-hit rate is low. Across the 100–200 evaluation budgets, the top-5 success rate is 0.9 and the top-10 success rate is 1.0, whereas the top-3 success rate remains between 0.3 and 0.4 and the exact-hit rate reaches only 0.1. Thus, the raw hardmax BO runs usually enter the near-optimal region of the cutout benchmark landscape, but they do not consistently distinguish the global optimum from the neighbouring high-ranked designs within this tight budget.

For pow0.5(hardmax), the top-k curves rise together with the exact-hit curve. The top-3 success rate is already 0.9 at 125 evaluations and reaches 1.0 by 175 evaluations, while the exact-hit rate reaches 1.0 at 200 evaluations. The GA curves show a slower movement into the best-ranked designs in the cutout benchmark landscape: GA has no exact hits, reaches top-3 success of only 0.1 from 150 evaluations onward, and reaches top-10 success of 0.5 at 200 evaluations. Because the power transform preserves the hardmax ordering, the GA top-k curves are identical for the two retained objectives. The target-ladder comparison therefore refines the reliability result: raw hardmax BO is already effective at finding near-optimal designs for the cutout benchmark, but the 0.5-power transform is what converts that near-optimal search behaviour into reliable exact-optimum recovery within 200 evaluations.

The first-hit curves in Fig. 13 use the same Kaplan–Meier estimator as Fig. 9, now applied to the 200-evaluation cutout runs. Runs that do not find the global optimum by the end of the budget are censored at 200 evaluations.

For the raw hardmax objective, the BO curve drops only once, at evaluation 142, because only one of the ten BO runs reaches the exact global optimum by the 200-evaluation budget. The other BO runs, like all GA runs, remain without an exact hit over this budget window. This agrees with the target-ladder evidence: raw hardmax BO often reaches near-optimal designs for the cutout benchmark, but exact rank-1 discovery is still rare in the tight-budget regime.

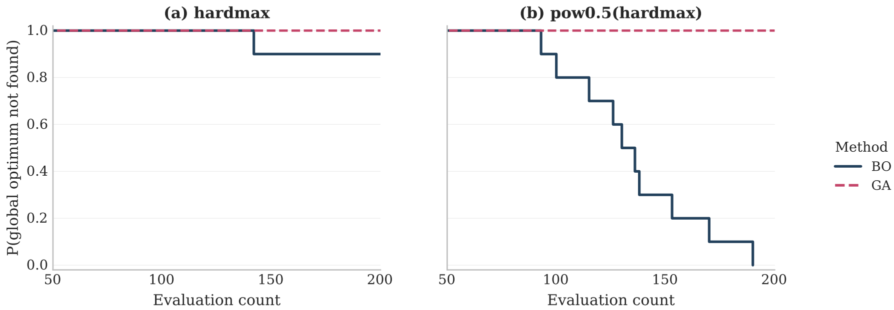


Fig. 13 Kaplan–Meier first-hit curves for the cutout benchmark at a 200-evaluation budget. The vertical axis gives the estimated probability, over the 10 independent runs, that a run has not yet found the global optimum by the corresponding evaluation count; lower curves therefore indicate earlier and more reliable discovery

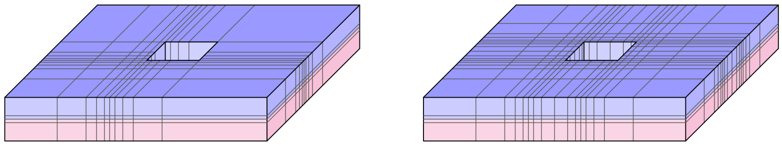
For pow0.5(hardmax), the BO curve drops steadily from the first hit at evaluation 93 to the final hit at evaluation 190, with median first-hit evaluation 133. The GA curve remains flat because no GA run reaches the exact optimum by 200 evaluations. Relative to raw hardmax BO and equal-budget GA, the transformed objective improves not only the final exact-hit reliability, but also the discovery dynamics: the BO converges to the global optimum earlier and more consistently within the same budget. In the context of a high-fidelity DG evaluator, this earlier first hit is important because it means that the best laminate for the cutout benchmark is identified after fewer objective queries within the same finite design library.

The matched-reliability comparison for the cutout benchmark is therefore focused on pow0.5(hardmax), the retained objective for which BO reaches complete reliability within the tight-budget window. Table 9 uses the 200-evaluation BO result as the reference case and reports the GA budget needed to reach the same 10/10 global-optimum success rate.

Table 9 Matched-reliability summary for the cutout benchmark using pow0.5(hardmax). The BO reference budget is 200 evaluations. Ratios are reported as GA/BO. First-hit and time values are means over successful runs only. Estimated time assumes 10 s per high-fidelity cutout objective evaluation plus stored optimiser/lookup overhead. The break-even cost is the DG evaluation time at which BO and the matched GA have the same estimated mean time to first hit.

Objective	BO target success	Budget			Mean first hit			Time to first hit [min]			Break-even (s/eval)
		BO	GA	GA/BO	BO	GA	GA/BO	BO	GA	GA/BO	
pow0.5(hardmax)	1.0	200	2880	14.40	135.1	1094.5	8.10	24.3	182.4	7.50	0.11

The table shows that matching the reliability already obtained by BO requires 2880 GA evaluations, or 14.4 times the BO reference budget. The same evaluation-cost gap appears in the first-hit statistics: even after the GA budget is increased until all runs succeed, the successful GA runs first reach the optimum after about eight times as many objective evaluations as BO. Using the same timing model as in the



(a) One-corner optimisation mesh **(b)** Four-corner validation mesh

Fig. 14 Cutout mesh pair used for optimisation-level validation: (a) the one-corner refined mesh used to construct the cutout objective landscapes, with 336 hexahedral elements, and (b) the four-corner refined mesh used as the validation mesh for selected designs, with 880 hexahedral elements

free-edge benchmark, but with the cutout DG evaluation cost set to $c_{\mathrm{DG}} = 10$ s, this gives an estimated mean time to first hit of 24.3 min for BO and 182.4 min for the matched GA. The break-even DG cost is only 0.11 s per evaluation, far below the cutout DG evaluation time used here. For the geometry-induced cutout problem, this is the practical value of the transformed BO objective: at the same 10/10 empirical reliability, BO identifies the best laminate about 7.5 times faster than the matched GA.

This comparison is deliberately conservative in one respect: it compares BO and GA while holding the DG evaluator fixed, so it captures only the optimisation-layer saving. A broader DG–BO versus conventional conforming three-dimensional finite-element GA comparison would also include the analysis-layer cost difference, because many more GA evaluations would be coupled to a more expensive low-aspect-ratio solid-mesh analysis. The earlier DG validation study measured an approximately 32-fold runtime reduction relative to a hierarchical conforming finite-element reference for a cross-ply free-edge benchmark [37]; an equivalent conforming finite-element timing study was not performed for the present cutout landscape, so this factor is not included in the reported speedups. As an illustrative scaling, if $k = c_{\mathrm{conf}}/c_{\mathrm{DG}}$ denotes the per-evaluation cost ratio for a hypothetical conforming finite-element replacement, the cutout matched-reliability timing ratio would scale approximately as $7.5k$. Even a modest $k = 5$ would imply an end-to-end ratio of about 38, while the $k \approx 32$ ratio measured in the earlier cross-ply benchmark would give an illustrative ratio of about 240. These values should be read only as scaling estimates, not as measured cutout conforming finite-element results.

The cutout benchmark objective landscapes are constructed with the one-corner refined mesh in Fig. 14(a), which concentrates refinement at the monitored cutout corner. This is the corner used to define the optimisation objective, but the physical cutout also contains other corners where local stress concentrations can develop. The validation question is whether refining those other corners changes the ordering of the best designs in the cutout benchmark landscape. To answer this targeted question, the leading designs for both objectives, hardmax and pow0.5(hardmax), were re-analysed on the four-corner refined mesh in Fig. 14(b), with the resulting rank changes and objective-value discrepancies reported in Table 10.

The top-ten ordering is unchanged under the four-corner validation mesh for both retained objectives. In particular, the rank-1 design used as the global optimum for the cutout benchmark remains the best design among the re-analysed top-ten candidates.

Table 10 Four-corner mesh validation for the top-ten cutout designs. The table compares the ranking and objective values of the top-ten designs from the one-corner optimisation mesh after re-analysis with the four-corner validation mesh, for hardmax and pow0.5(hardmax). The rank change is $\Delta r = r_{\text{four-corner}} - r_{\text{one-corner}}$, after re-ranking the same ten designs with the four-corner mesh. The objective discrepancy is $e_J = 100(J_{\text{one-corner}} - J_{\text{four-corner}})/J_{\text{four-corner}}$.

Objective	Checked ranks	Unchanged ranks	Mean $\lvert\Delta r\rvert$	Max $\lvert\Delta r\rvert$	Mean $\lvert e_J\rvert$ [%]	Max $\lvert e_J\rvert$ [%]
hardmax	1–10	10/10	0.0	0	1.25	1.57
pow0.5(hardmax)	1–10	10/10	0.0	0	0.62	0.78

The four-corner mesh gives slightly lower objective values for these candidates, but the discrepancies are small: the mean absolute error is 1.25% for hardmax and 0.62% for pow0.5(hardmax), with corresponding maxima of 1.57% and 0.78%. This does not claim that the one-corner-refined mesh reproduces the complete four-corner-refined landscape for every possible laminate. It supports the more targeted decision needed for the benchmark: the near-optimal designs for the cutout benchmark, including the selected optimum, are stable under the more refined validation mesh.

The two benchmarks can now be compared as finite objective landscapes for the same design library and the same retained transformed objective. Table 11 summarises this comparison for pow0.5(hardmax). The top-ranked sets are distinct: none of the top-five or top-ten designs in the free-edge benchmark landscape appears in the corresponding top-ranked set of the cutout benchmark landscape. The rank-1 designs are also problem-specific. The free-edge benchmark optimum, [IM7/MTM45-1@90° / IM6/3501-6@90°], is ranked 1329th in the cutout benchmark landscape, while the cutout benchmark optimum, [T800/924C@45° / T700/MTM57@45°], is ranked 827th in the free-edge benchmark landscape.

Table 11 Cross-benchmark comparison for pow0.5(hardmax). Ranks are computed within each complete 11556-design objective landscape, with rank 1 denoting the benchmark-specific global optimum.

Comparison quantity	Result
Shared designs in the two top-five sets	0/5
Shared designs in the two top-ten sets	0/10
Free-edge optimum rank in the cutout benchmark landscape	1329/11556
Cutout optimum rank in the free-edge landscape	827/11556

This low overlap is useful for interpreting the role of the second benchmark. The cutout case is not merely a repeat of the free-edge optimisation with the same preferred laminate; it changes the stress mechanism enough that the best material-angle combination changes with it. In the free-edge case, the leading designs are dominated by 90° groups associated with the straight-edge interlaminar-stress response. In the cutout case, the best design uses 45° groups around a local geometric discontinuity. The combined results therefore support the central purpose of the benchmark pair: the

DG-BO framework can search the same finite laminate library under different three-dimensional stress mechanisms, while the objective transformation that improved BO reliability in the free-edge study remains effective in the geometry-induced cutout problem.

4 Conclusions

This study has shown that local interlaminar-stress optimisation can be treated as a direct high-fidelity laminate design problem rather than as a post-design check. The proposed framework couples a high-aspect-ratio full 3D DG finite-element evaluator with BO over a finite material-and-fibre-angle laminate set. By permitting local refinement with stretched hexahedral elements, the DG evaluator resolves the monitored interlaminar-stress region without requiring the surrounding domain to be filled with the small low-aspect-ratio elements typically needed in conventional 3D solid meshes. The DG component then supplies resolved interlaminar normal and shear stresses over monitored patches, while BO uses the accumulated objective data to decide which discrete laminate should receive the next high-fidelity evaluation. This pairing is important: the analysis model makes the local stress objective available, and the optimiser keeps the number of such analyses small enough for repeated search.

The results demonstrate that the numerical form of the scalar objective matters, even when all variants are built from the same DG stress data and the same candidate laminate set. In the free-edge benchmark, the best-performing transformed objectives allowed BO to find the global optimum in every run across the tested budgets, whereas the equal-budget GA found it in only 20% of runs. After the GA budget was increased to match this reliability, GA required about seven times more evaluations on average to find the same optimum, corresponding to about 6.7 times longer mean time to first hit. In the cutout benchmark, where the stress field is driven by a geometry-induced corner concentration, the best transformed objective again led BO to find the optimum in every run, while the equal-budget GA found no exact optima. The mean BO first hit occurred after 135.1 evaluations, about 1.2% of the feasible design space, while the matched GA required about eight times more evaluations and about 7.5 times longer mean time to identify it.

Importantly, these timing ratios compare BO and GA while holding the DG evaluator fixed, so they capture only the optimisation-layer benefit. The broader DG–BO framework is expected to be more advantageous against a conventional conforming three-dimensional finite-element GA workflow, where many more GA evaluations would be coupled to a more expensive low-aspect-ratio solid-mesh analysis. In that sense, the reported timing gains are conservative: they do not include the additional per-evaluation saving enabled by the high-aspect-ratio DG solver.

Overall, the study supports DG–BO as a practical route for directly optimising resolved local interlaminar-stress objectives under limited high-fidelity evaluation budgets. By combining an efficient full 3D DG stress evaluator with acquisition-driven BO, the framework turns a response quantity that is usually treated indirectly or after design into a searchable optimisation objective.

Appendix A Free-edge benchmark raw objective results

Table 12 Raw free-edge optimisation results for all objective variants and evaluation budgets. SR denotes the global-optimum success rate over ten independent runs; MFH denotes the mean first-hit evaluation over successful runs only. Objective labels use compact mathematical notation: H is hardmax, $S2$, $S4$ and $S6$ are smoothmax objectives with $q = 2$, 4 and 6, superscripts denote power transforms, and $\log H$ denotes a logarithmic transform of H.

	550				575				600				625				650			
	BO		GA		BO		GA		BO		GA		BO		GA		BO		GA	
Objective	SR	MFH	SR	MFH	SR	MFH	SR	MFH	SR	MFH	SR	MFH	SR	MFH	SR	MFH	SR	MFH	SR	MFH
H	0.8	172.0	0.2	335.5	1.0	305.5	0.2	335.5	1.0	352.2	0.2	335.5	1.0	287.9	0.2	335.5	1.0	309.4	0.2	335.5
$S6$	0.8	200.6	0.1	521.0	1.0	293.8	0.1	521.0	1.0	330.0	0.1	521.0	1.0	297.3	0.1	521.0	1.0	310.5	0.1	521.0
$S4$	0.8	193.6	0.2	249.5	0.9	247.7	0.2	249.5	1.0	302.0	0.2	249.5	1.0	294.8	0.2	249.5	1.0	304.9	0.2	249.5
$S2$	1.0	229.3	0.1	61.0	0.9	257.7	0.1	61.0	1.0	319.0	0.1	61.0	1.0	294.4	0.1	61.0	1.0	282.4	0.2	347.0
$\log H$	0.7	331.9	0.2	335.5	0.8	328.1	0.2	335.5	0.8	424.1	0.2	335.5	0.8	340.8	0.2	335.5	0.8	380.9	0.2	335.5
$H^{0.3}$	1.0	154.9	0.2	335.5	1.0	225.8	0.2	335.5	1.0	285.8	0.2	335.5	1.0	286.1	0.2	335.5	1.0	354.5	0.2	335.5
$H^{0.5}$	1.0	215.0	0.2	335.5	1.0	208.9	0.2	335.5	1.0	207.1	0.2	335.5	1.0	251.1	0.2	335.5	1.0	300.6	0.2	335.5
$H^{0.7}$	0.9	187.7	0.2	335.5	1.0	218.6	0.2	335.5	1.0	268.0	0.2	335.5	1.0	286.1	0.2	335.5	0.9	282.0	0.2	335.5
$S6^{0.3}$	1.0	184.0	0.1	521.0	1.0	242.3	0.1	521.0	1.0	255.9	0.1	521.0	1.0	281.2	0.1	521.0	1.0	308.9	0.1	521.0
$S6^{0.5}$	1.0	167.3	0.1	521.0	1.0	232.1	0.1	521.0	1.0	192.1	0.1	521.0	1.0	236.7	0.1	521.0	1.0	293.0	0.1	521.0
$S6^{0.7}$	0.8	186.1	0.1	521.0	1.0	205.9	0.1	521.0	1.0	255.9	0.1	521.0	1.0	255.9	0.1	521.0	0.9	351.4	0.1	521.0
$S4^{0.3}$	1.0	241.1	0.2	249.5	1.0	223.0	0.2	249.5	1.0	243.5	0.2	249.5	1.0	283.5	0.2	249.5	1.0	367.0	0.2	249.5
$S4^{0.5}$	1.0	196.4	0.2	249.5	1.0	257.5	0.2	249.5	1.0	208.0	0.2	249.5	1.0	249.9	0.2	249.5	1.0	299.2	0.2	249.5
$S4^{0.7}$	0.9	190.3	0.2	249.5	1.0	242.9	0.2	249.5	1.0	202.6	0.2	249.5	1.0	233.6	0.2	249.5	1.0	250.0	0.2	249.5
$S2^{0.3}$	1.0	175.8	0.1	61.0	1.0	219.4	0.1	61.0	1.0	255.0	0.1	61.0	1.0	254.5	0.1	61.0	1.0	303.9	0.2	347.0
$S2^{0.5}$	1.0	191.3	0.1	61.0	1.0	195.7	0.1	61.0	1.0	233.2	0.1	61.0	1.0	219.6	0.1	61.0	1.0	263.0	0.2	347.0
$S2^{0.7}$	1.0	184.3	0.1	61.0	0.9	267.6	0.1	61.0	1.0	237.3	0.1	61.0	1.0	232.3	0.1	61.0	1.0	308.3	0.2	347.0

Appendix B GA hyperparameter sensitivity

The GA results in the main text use the settings in Table 4: population size $N_p = 20$, per-gene mutation probability $p_m = 1/d$ with $d = 4$, and tournament size $\tau = 3$. To check that the BO–GA conclusions are not an artefact of this single GA setting, a small hardmax-only sensitivity study was performed. The sensitivity grid tested $N_p \in \{10, 20, 50\}$, $p_m \in \{1/d, 2/d\}$ and $\tau \in \{2, 3\}$ for both benchmarks, using the same ten seeds and the same total-evaluation budgets as the hardmax GA runs in the main text. All 720 sensitivity runs used the exact requested evaluation budget; partial final offspring batches were allowed by the existing GA implementation when required.

Figure 15 compares the exact global-optimum success rate for all tested GA settings at each benchmark budget. Changing the GA hyperparameters produces mixed effects: some settings modestly improve exact recovery, some leave it unchanged, and others reduce it. In the free-edge benchmark, the best tested variants increase exact hits from 2/10 for the paper setting to at most 3/10 at 550 and 600 evaluations and 4/10 at 650 evaluations. In the cutout benchmark, exact global-optimum recovery remains at most 1/10 for every tested budget.

Appendix C Analytical random-search reference

The BO and GA results in the main text are empirical optimiser outcomes: each method uses its own search rule, and success is measured over repeated seeded runs. A blind random-search reference can be treated more directly because the candidate library is finite and the sampling process is known. Consider uniform random search without replacement over the same feasible laminate library $\mathcal{X}_{\text{feas}}$. This is the relevant

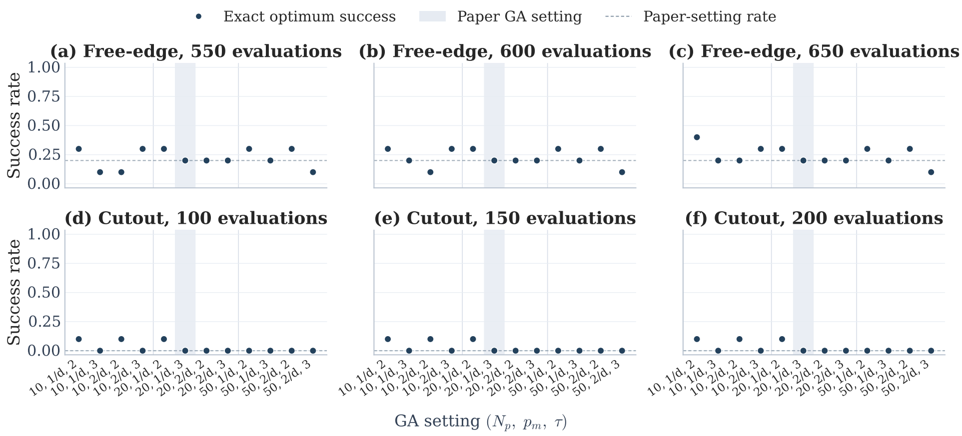


Fig. 15 GA hyperparameter sensitivity for the hardmax objective. Each panel reports the exact global-optimum success rate ($r = 1$) over ten independent runs for all tested settings (N_p, p_m, τ). The shaded band marks the paper GA setting, $N_p = 20$, $p_m = 1/d$ and $\tau = 3$; the dashed horizontal line shows the corresponding paper-setting success rate in that panel

random-search convention here, because evaluating the same laminate twice would not add information and both BO and GA exclude or repair duplicate objective queries. If $N = |\mathcal{X}_{\text{feas}}|$, $s = |\mathcal{X}_J^\star|$ and B distinct designs are sampled, the probability of hitting at least one globally optimal design for objective J is

$$P_{\text{RS}}^\star(B) = 1 - \frac{\binom{N-s}{B}}{\binom{N}{B}}, \qquad B \le N. \tag{C1}$$

Thus, for exact global-optimum discovery, the objective landscape affects random search only through the number of tied global optima s. In the benchmark landscapes used here, the exact optimum is unique for all reported free-edge objective variants and for both retained cutout objectives, so $s = 1$. Because both benchmarks also use the same feasible laminate library, $N = 11556$, Eq. (C1) reduces in all these cases to

$$P_{\text{RS}}^\star(B) = \frac{B}{N}. \tag{C2}$$

Table 13 gives this analytical reference for the two budget ranges used in the main benchmark studies. The comparison is intentionally limited to exact rank-1 discovery: it asks how likely blind uniform sampling would be to stumble on the same unique optimum targeted by BO and GA. The answer is small across both studies. At the largest free-edge budget, 650 evaluations sample only 5.63% of the library; at the largest cutout budget, 200 evaluations sample only 1.73%. These values provide a closed-form sanity check on the main results: the complete BO exact-hit outcomes reported for the strongest transformed objectives are not explained by the budgets being large enough for blind sampling to be likely to find the optimum.

Table 13 Analytical random-search reference for exact global-optimum discovery. Random search is uniform without replacement over the 11,556-design feasible library. Since all reported benchmark objective landscapes have a unique global optimum, the probability depends only on the evaluation budget.

Benchmark budget range	Budget B	Sampled library [%]	$P^{\star}_{\mathrm{RS}}(B)$ [%]
Free edge	550	4.76	4.76
Free edge	575	4.98	4.98
Free edge	600	5.19	5.19
Free edge	625	5.41	5.41
Free edge	650	5.63	5.63
Cutout	100	0.87	0.87
Cutout	125	1.08	1.08
Cutout	150	1.30	1.30
Cutout	175	1.51	1.51
Cutout	200	1.73	1.73

Appendix D Post-hoc landscape interpretation

This appendix reports post-hoc diagnostics used to interpret the enumerated hardmax objective landscapes. The diagnostics were not optimisation models and were not used to guide BO or GA; they were applied only after the DG objective values had been computed for every feasible design. Their purpose is to identify which material-angle features are associated with the ranked landscape structure, and to support the physical interpretation of the benchmark optima. For each benchmark, the analysed response was

$$y = \log_{10} J_{\mathrm{hard}}, \tag{D1}$$

which reduces the dynamic range of the objective values while preserving their ordering. Each design was encoded by angle-structure features and by material mismatch features. The angle features were the same-angle indicator, the absolute angle difference and a set of binary indicators for the ordered angle pair, so that, for example, 45°/90° and 90°/45° are treated as distinct categories. For a material property p, the between-group mismatch was measured as

$$m_p = \log\left(\frac{\max(p_1, p_2)}{\min(p_1, p_2)}\right), \tag{D2}$$

where p_1 and p_2 are the two group properties. Thus $m_p = 0$ corresponds to matched material values, and larger values indicate a stronger mismatch. The material properties included E_1, E_2, E_3, G_{12}, G_{23}, G_{31}, ν_{12}, ν_{23} and ν_{13}.

The global diagnostic used an ExtraTrees regression model [19] to predict y from these encoded features. The model was used only as an interpretive device: if a feature group is important, randomly permuting that group in held-out data should reduce predictive accuracy. For a feature group g, the grouped permutation score was computed as

$$I_g = R^2\left(\hat{f}(X), y\right) - R^2\left(\hat{f}(X_{\pi(g)}), y\right), \tag{D3}$$

where $X_{\pi(g)}$ denotes the held-out feature matrix after the columns in group g have been jointly permuted. The reported percentages in Table 14 normalise the positive I_g values so that they sum to 100%. In both benchmarks the model used 800 trees, a 25% held-out test split, 40 permutation repeats and a fixed random seed of 42.

Table 14 Summary of the post-hoc tree-ensemble diagnostics for the hardmax landscapes. The importance percentages are grouped permutation importances normalised over positive R^2 drops.

Benchmark	Designs	Test R^2	Same angle	Angle pair	$\lvert\Delta\theta\rvert$	Material ratios
Free edge	11,556	0.982	54.6%	31.6%	8.0%	5.8%
Cutout	11,556	0.972	53.7%	24.5%	21.3%	0.5%

The tree model identifies broad landscape structure, but it does not by itself explain which material contrasts matter inside a fixed angle family. For that reason, a second diagnostic held the angle family fixed and computed Spearman rank correlations between each mismatch feature and y,

$$\rho_s(p) = \mathrm{corr}(\mathrm{rank}(m_p), \mathrm{rank}(y))\,. \tag{D4}$$

This step asks a narrower question: once the angle pair has been chosen, is there a simple monotonic relationship between a material mismatch and the objective? Table 15 reports the strongest correlations for the angle families discussed in the main text.

Table 15 Stratified Spearman diagnostics within selected same-angle families. Positive ρ_s means that larger mismatch is associated with a larger objective; negative ρ_s means that larger mismatch is associated with a smaller objective within that fixed angle family.

Benchmark	Angle family	Strongest feature(s)	ρ_s	Interpretation
Free edge	90°/90°	ν_{12}, ν_{13}	0.986	The objective rises strongly when the two groups differ in either Poisson ratio; the global optimum has both values matched.
Free edge	0°/0°	ν_{12}, ν_{13}	0.672	The same between-group Poisson-ratio effect is visible, but less strongly than in the 90°/90° family.
Free edge	45°/45°	G_{12}, G_{31}	0.666	The clearest material effect shifts to shear: larger contrasts in G_{12} or G_{31} are associated with higher objectives.
Free edge	20°/20°	G_{12}, G_{31}	0.601	A similar shear-modulus association persists at this lower off-axis angle, although slightly more weakly.
Cutout	45°/45°	ν_{23}	−0.225	The largest trend is weak-to-moderate and negative; no single material mismatch explains this leading cutout family.
Cutout	20°/20°	ν_{23}	−0.219	The same weak negative ν_{23} trend appears, again indicating no dominant one-ratio rule.

Finally, for each ordered material assignment, the objective values were compared over the available angle pairs and the best angle pair was recorded. There are 144 ordered material assignments in each benchmark, corresponding to 12 possible materials in the first group and 12 in the second. Same-angle choices are infeasible for the 12 same-material assignments, because identical material-angle groups are excluded from the design library; for the remaining 132 assignments, a same-angle pair gives the lowest objective in every case for both benchmarks. Table 16 summarises this fixed-material result, while Figs. 16 and 17 show the corresponding material-by-material maps and mark the global optimum in each landscape. Together, they also show that the preferred same-angle family changes with the benchmark geometry. This is the basis for the interpretation in the main text: the global optima are not explained by one universally preferred angle or one universally preferred material ratio, but by geometry-dependent interactions between angle family and material compatibility.

Table 16 Best angle-pair counts when each ordered material assignment is held fixed and only the angle pair is varied. Same-angle choices are feasible for 132 of the 144 ordered material assignments; the remaining 12 are same-material assignments for which identical material-angle groups are excluded.

Benchmark	Feasible same-angle winners	Most frequent winning angle pairs
Free edge	132/132	$0°/0°$ (80), $90°/90°$ (26), $-45°/-45°$ (15), $-20°/20°$ (10)
Cutout	132/132	$45°/45°$ (81), $20°/20°$ (45), $20°/0°$ (12), $0°/0°$ (6)

For the free-edge benchmark, these diagnostics provide the full-landscape support for the interpretation used in Section 3.1. The hardmax ranking is strongly concentrated in same-angle designs: the first 663 designs are all same-angle pairs, and the first mixed-angle design appears only at rank 664. At the same time, the fixed-material sweep, reported in Table 16 and Fig. 16, shows that the all-90° top-three set should not be generalised into a universal 90° rule. The most frequent fixed-material winner is $0°/0°$ (80 cases), whereas $90°/90°$ wins 26 cases. Within fixed same-angle families, the material contrast also changes with the angle family: Table 15 shows that the $90°/90°$ family is most strongly associated with between-group mismatches in ν_{12} and ν_{13}, while the $45°/45°$ family is most strongly associated with between-group mismatches in G_{12} and G_{31}. The free-edge hardmax global optimum lies at the favourable end of the $90°/90°$ trend, with matched ν_{12} and matched ν_{13} between the two materials, so both corresponding property ratios are 1. Thus the free-edge optimum is best interpreted as an interaction between the straight-edge loading mechanism, the selected same-angle family and material compatibility.

For the cutout benchmark, the same diagnostics show that the change from a straight free edge to a corner discontinuity changes the preferred angle family. The tree-ensemble diagnostic again identifies angle structure as the dominant source of explanatory information, with the same-angle indicator, ordered angle pair and absolute angle difference accounting for most of the positive grouped permutation importance in Table 14. The ranked hardmax landscape gives the same message directly: same-angle designs form only 10.3% of the feasible library, but the first 333 ranked designs are all same-angle pairs, and the first mixed-angle design appears only at rank 334. The fixed-material sweep also points to a same-angle preference, with same-angle pairs winning all 132 feasible ordered material assignments. Within those fixed-material winners, however, the dominant families are off-axis: either $45°/45°$

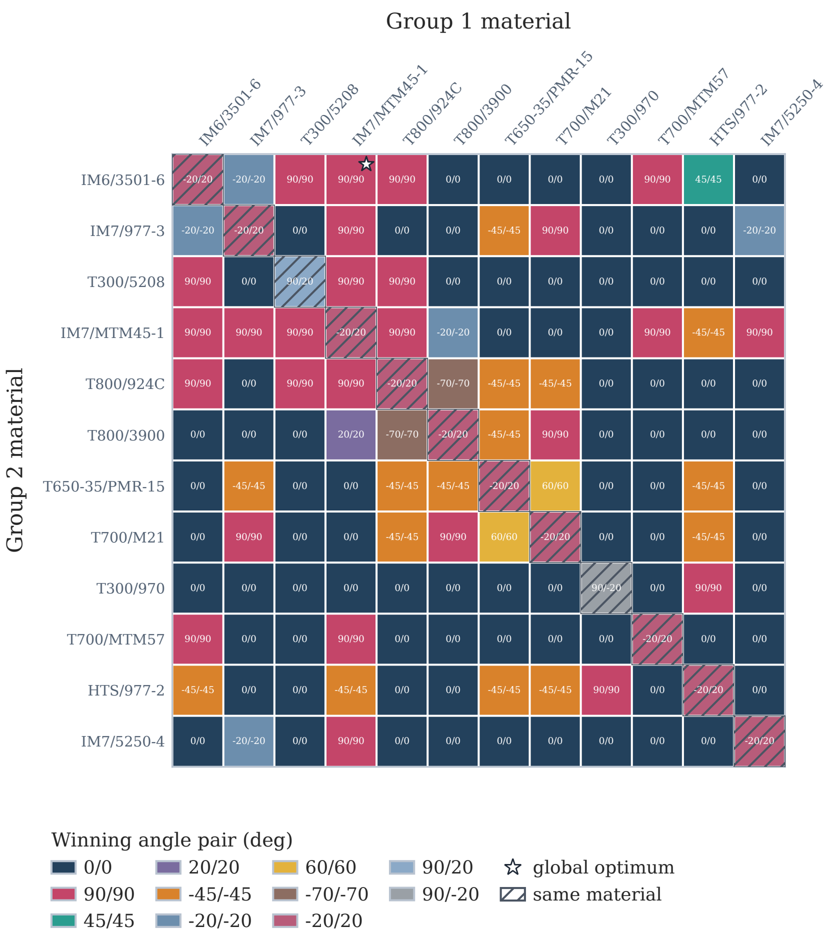


Fig. 16 Fixed-material winning angle-pair matrix for the free-edge hardmax landscape. Columns denote the group-1 material and rows denote the group-2 material. Each cell reports the angle pair that gives the lowest objective value for that ordered material assignment. Hatched diagonal cells are same-material assignments, for which same-angle choices are infeasible because identical material-angle groups are excluded from the design library

or $20°/20°$ gives the lowest cutout objective for 126 of the 132 feasible assignments. Unlike the free-edge $90°/90°$ family, the stratified Spearman checks for the leading cutout families do not show a strong single material-ratio rule; the largest monotonic trends are associated with ν_{23} mismatch and are only weak to moderate. The cutout optimum is therefore best interpreted as a geometry-dependent same-angle, off-axis family effect [47].

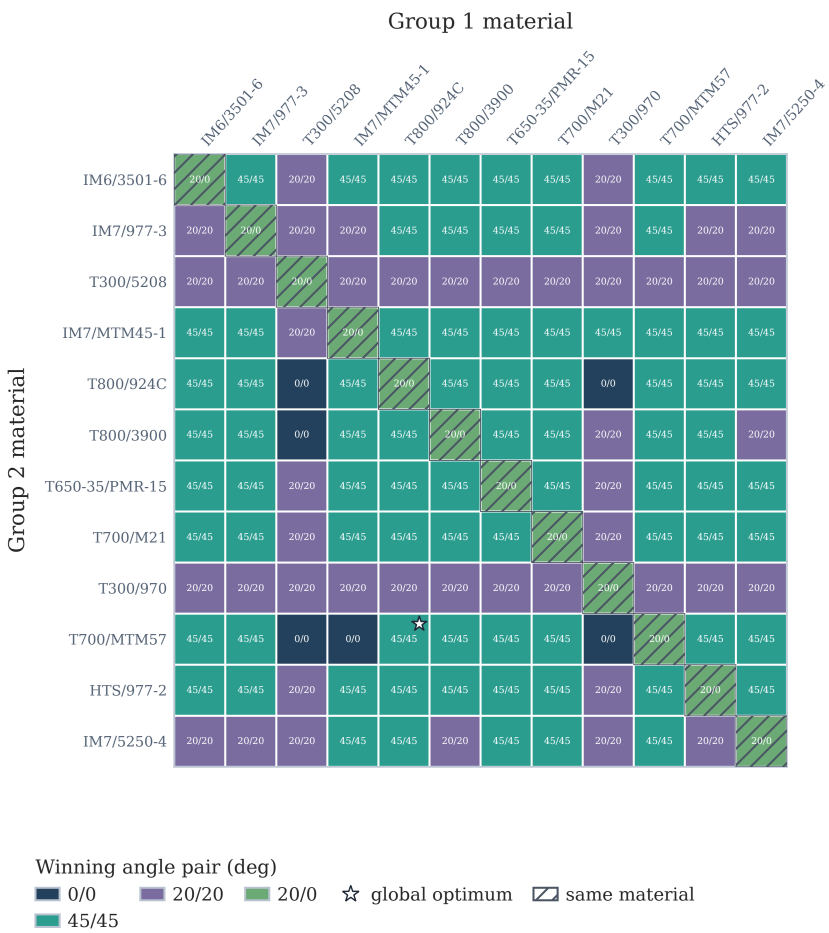


Fig. 17 Fixed-material winning angle-pair matrix for the cutout hardmax landscape. Columns denote the group-1 material and rows denote the group-2 material. Each cell reports the angle pair that gives the lowest objective value for that ordered material assignment. Hatched diagonal cells are same-material assignments, for which same-angle choices are infeasible because identical material-angle groups are excluded from the design library

Statements and Declarations

Funding

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) through a doctoral studentship awarded to Soheil Navvabi (project reference 2890415) within the EPSRC and NERC Centre for Doctoral Training in Offshore Wind Energy and the Environment (Aura CDT; grant EP/S023763/1).

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Author contributions

Soheil Navvabi: Conceptualization, Methodology, Validation, Software, Visualization, Writing – original draft, Writing – review & editing. **Stefano Giani:** Conceptualization, Methodology, Validation, Supervision, Writing – review & editing, Funding acquisition. **William M. Coombs:** Conceptualization, Methodology, Validation, Supervision, Writing – review & editing, Funding acquisition.

Data availability

The datasets generated and analysed during this study are openly available in the Durham Research Online DATAsets Archive under the Creative Commons Attribution 4.0 International licence at https://doi.org/10.15128/r2g445cd22w [36].

Use of artificial intelligence tools

During the preparation of this work, the authors used OpenAI ChatGPT to assist with editing, language polishing and consistency checks. The authors reviewed and revised the content as necessary and take full responsibility for the final manuscript.

References

[1] Albazzan MA, Harik R, Tatting BF, et al (2019) Efficient design optimization of nonconventional laminated composites using lamination parameters: A state of the art. Compos Struct 209:362–374. https://doi.org/10.1016/j.compstruct.2018.10.095

[2] Almeida JHS Jr., Ribeiro ML, Tita V, et al (2017) Stacking sequence optimization in composite tubes under internal pressure based on genetic algorithm accounting for progressive damage. Compos Struct 178:20–26. https://doi.org/10.1016/j.compstruct.2017.07.054

[3] Ament S, Daulton S, Eriksson D, et al (2023) Unexpected improvements to expected improvement for Bayesian optimization. In: Oh A, Naumann T, Globerson A, et al (eds) Advances in Neural Information Processing Systems, vol 36. Curran Associates, Inc., pp 20577–20612, https://doi.org/10.52202/075280-0904, URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/419f72cbd568ad62183f8132a3605a2a-Abstract-Conference.html, arXiv:2310.20708

[4] An H, Chen S, Huang H (2018) Multi-objective optimization of a composite stiffened panel for hybrid design of stiffener layout and laminate stacking sequence. Struct Multidiscip Optim 57:1411–1426. https://doi.org/10.1007/s00158-018-1918-2

[5] Balandat M, Karrer B, Jiang DR, et al (2020) BoTorch: A framework for efficient Monte-Carlo Bayesian optimization. In: Larochelle H, Ranzato M, Hadsell R, et al (eds) Advances in Neural Information Processing Systems, vol 33. Curran Associates, Inc., pp 21524–21538, URL https://proceedings.neurips.cc/paper/2020/hash/f5b1b89d98b7286673128a5fb112cb9a-Abstract.html, arXiv:1910.06403

[6] Beasley D, Bull DR, Martin RR (1993) An overview of genetic algorithms: Part 1, fundamentals. Univ Comput 15(2):58–69. URL https://orca.cardiff.ac.uk/id/eprint/64436

[7] Burhan M, Scalici T, Ullah Z, et al (2024) Investigating factors influencing interlaminar stresses and energy release rates of semi-elliptical cracks at free edges. Engineering Fracture Mechanics 307:110274. https://doi.org/10.1016/j.engfracmech.2024.110274

[8] Callahan KJ, Weeks GE (1992) Optimum design of composite laminates using genetic algorithms. Compos Eng 2(3):149–160. https://doi.org/10.1016/0961-9526(92)90001-m

[9] Chen S, Lin Z, An H, et al (2013) Stacking sequence optimization with genetic algorithm using a two-level approximation. Struct Multidiscip Optim 48:795–805. https://doi.org/10.1007/s00158-013-0927-4

[10] Chen Y, Xiao Z, Yang Y, et al (2025) Multi-objective optimization for impact resistance of composite laminates with non-conventional ply orientations: An integrated finite element and machine learning framework. Thin-Walled Struct 216:113687. https://doi.org/10.1016/j.tws.2025.113687

[11] Chuaqui TRC, Rhead AT, Butler R, et al (2021) A data-driven Bayesian optimisation framework for the design and stacking sequence selection of increased notched strength laminates. Compos Part B Eng 226:109347. https://doi.org/10.1016/j.compositesb.2021.109347

[12] Cowen-Rivers AI, Lyu W, Tutunov R, et al (2022) HEBO: Pushing the limits of sample-efficient hyperparameter optimisation. Journal of Artificial Intelligence Research 74:1269–1349. https://doi.org/10.1613/jair.1.13643

[13] Damghani M, Matthews J, Murphy A, et al (2023) Numerical shape, thickness and stacking sequence optimisation and experimental study of hybrid composite plates under in-plane shear loading. Structures 51:372–390. https://doi.org/10.1016/j.istruc.2023.03.066

[14] Fedon N, Weaver PM, Pirrera A, et al (2021) A method using beam search to design the lay-ups of composite laminates with many plies. Compos Part C Open Access 4:100072. https://doi.org/10.1016/j.jcomc.2020.100072

[15] Fernández F (2026) Optimal design strategy for variable-stiffness composite laminates addressing strength and manufacturing challenges. Optim Eng 27:261–316. https://doi.org/10.1007/s11081-025-10013-0

[16] Frazier PI (2018) Bayesian Optimization. In: Gel E, Ntaimo L (eds) Recent Advances in Optimization and Modeling of Contemporary Problems. INFORMS TutORials in Operations Research, INFORMS, Catonsville, MD, p 255–278, https://doi.org/10.1287/educ.2018.0188

[17] Garcia-Pis A, Gogu C, Michel L, et al (2025) A unified mixed-variables framework for manufacturing-constrained optimization of composite laminates applied to conventional and Double-Double layups. Struct Multidiscip Optim 68:176. https://doi.org/10.1007/s00158-025-04101-z

[18] Garrido-Merchán EC, Hernández-Lobato D (2020) Dealing with categorical and integer-valued variables in Bayesian optimization with Gaussian processes. Neurocomputing 380:20–35. https://doi.org/10.1016/j.neucom.2019.11.004, arXiv:1805.03463 [stat.ML]

[19] Geurts P, Ernst D, Wehenkel L (2006) Extremely randomized trees. Machine Learning 63:3–42. https://doi.org/10.1007/s10994-006-6226-1

[20] Ghiasi H, Pasini D, Lessard L (2009) Optimum stacking sequence design of composite materials Part I: Constant stiffness design. Compos Struct 90(1):1–11. https://doi.org/10.1016/j.compstruct.2009.01.006

[21] Ghiasi H, Fayazbakhsh K, Pasini D, et al (2010) Optimum stacking sequence design of composite materials Part II: Variable stiffness design. Compos Struct 93(1):1–13. https://doi.org/10.1016/j.compstruct.2010.06.001

[22] Hermansen SM, Lund E (2023) Multi-material and thickness optimization of laminated composite structures subject to high-cycle fatigue. Struct Multidiscip Optim 66(12):259. https://doi.org/10.1007/s00158-023-03708-4

[23] Irisarri FX, Lasseigne A, Leroy FH, et al (2014) Optimal design of laminated composite structures with ply drops using stacking sequence tables. Compos Struct 107:559–569. https://doi.org/10.1016/j.compstruct.2013.08.030

[24] Kalita K, Haldar S, Chakraborty S (2022) A comprehensive review on high-fidelity and metamodel-based optimization of composite laminates. Arch Comput Methods Eng 29:3305–3340. https://doi.org/10.1007/s11831-021-09699-z

[25] Kharghani N, Mittelstedt C (2022) Reduction of free-edge effects around a hole of a composite plate using a numerical layup optimization. Composite Structures 284:115139. https://doi.org/10.1016/j.compstruct.2021.115139

[26] Lakshmi K, Rao ARM (2015) Optimal design of laminate composite plates using dynamic hybrid adaptive harmony search algorithm. J Reinf Plast Compos 34(6):493–518. https://doi.org/10.1177/0731684415574228

[27] Lin CC, Lee YJ (2004) Stacking sequence optimization of laminated composite structures using genetic algorithm with local improvement. Compos Struct 63(3–4):339–345. https://doi.org/10.1016/S0263-8223(03)00182-X

[28] Liu X, Qin J, Zhao K, et al (2023) Design optimization of laminated composite structures using artificial neural network and genetic algorithm. Compos Struct 305:116500. https://doi.org/10.1016/j.compstruct.2022.116500

[29] Macquart T, Bordogna MT, Lancelot P, et al (2016) Derivation and application of blending constraints in lamination parameter space for composite optimisation. Compos Struct 135:224–235. https://doi.org/10.1016/j.compstruct.2015.09.016

[30] Mitchell M (1996) Implementing a genetic algorithm. In: An Introduction to Genetic Algorithms. The MIT Press, Cambridge, MA, p 151–175, https://doi.org/10.7551/mitpress/3927.003.0007

[31] Mittelstedt C, Becker W (2004) Interlaminar stress concentrations in layered structures: Part I – A selective literature survey on the free-edge effect since 1967. Journal of Composite Materials 38(12):1037–1062. https://doi.org/10.1177/0021998304040566

[32] Mittelstedt C, Becker W (2007) Free-edge effects in composite laminates. Applied Mechanics Reviews 60(5):217–245. https://doi.org/10.1115/1.2777169

[33] Montemurro M, Pagani A, Fiordilino GA, et al (2018) A general multi-scale two-level optimisation strategy for designing composite stiffened panels. Compos Struct 201:968–979. https://doi.org/10.1016/j.compstruct.2018.06.119

[34] Montemurro M, Izzi MI, El-Yagoubi J, et al (2019) Least-weight composite plates with unconventional stacking sequences: Design, analysis and experiments. J Compos Mater 53(16):2209–2227. https://doi.org/10.1177/0021998318824783

[35] Muc A (2016) Choice of design variables in the stacking sequence optimization for laminated structures. Mech Compos Mater 52(2):211–224. https://doi.org/10.1007/s11029-016-9574-3

[36] Navvabi S, Giani S, Coombs W (2026) Bayesian optimisation of resolved interlaminar stresses in composite laminates using high-aspect-ratio DG finite elements [dataset]. Durham Research Online DATAsets Archive, https://doi.org/10.15128/r2g445cd22w

[37] Navvabi S, Giani S, Coombs W (2026) Efficient 3D analysis of composite laminates using high-aspect-ratio discontinuous Galerkin finite elements. Composites

Communications 64:102830. https://doi.org/10.1016/j.coco.2026.102830

[38] Nguyen PD, Vu QV, Papazafeiropoulos G, et al (2024) A comprehensive investigation on optimum layer thickness and fiber orientations of laminated composite plates against buckling failure under various loading conditions. J Mech Sci Technol 38(9):4757–4763. https://doi.org/10.1007/s12206-024-0812-7

[39] Ntourmas G, Özcan E, Chronopoulos D, et al (2019) Towards a streamlined stacking sequence optimisation methodology for blended composite aircraft structures. In: Proceedings of the 8th European Conference for Aeronautics and Aerospace Sciences (EUCASS). EUCASS Association, https://doi.org/10.13009/EUCASS2019-632

[40] Ntourmas G, Glock F, Deinert S, et al (2023) Stacking sequence optimisation of an aircraft wing skin. Struct Multidiscip Optim 66:31. https://doi.org/10.1007/s00158-022-03483-8

[41] Peng X, Chen K, Jia W, et al (2025) Stacking sequence optimization of variable thickness composite laminated plate based on multi-peak stacking sequence table. Compos Struct 356:118886. https://doi.org/10.1016/j.compstruct.2025.118886

[42] Picchi Scardaoni M, Montemurro M (2020) A general global-local modelling framework for the deterministic optimisation of composite structures. Struct Multidiscip Optim 62:1927–1949. https://doi.org/10.1007/s00158-020-02586-4

[43] Picelli R, Townsend S, Brampton C, et al (2018) Stress-based shape and topology optimization with the level set method. Comput Methods Appl Mech Eng 329:1–23. https://doi.org/10.1016/j.cma.2017.09.001

[44] Pipes RB, Pagano NJ (1970) Interlaminar stresses in composite laminates under uniform axial extension. Journal of Composite Materials 4(4):538–548. https://doi.org/10.1177/002199837000400409

[45] Raju IS, Crews JH Jr. (1980) Interlaminar stress singularities at a straight free edge in composite laminates. Tech. Rep. NASA-TM-81876, NASA Langley Research Center, URL https://ntrs.nasa.gov/citations/19800024246

[46] Rasmussen CE, Williams CKI (2006) Gaussian Processes for Machine Learning. Adaptive Computation and Machine Learning, MIT Press, Cambridge, MA, https://doi.org/10.7551/mitpress/3206.001.0001

[47] Rybicki EF, Schmueser DW (1978) Effect of stacking sequence and lay-up angle on free edge stresses around a hole in a laminated plate under tension. Journal of Composite Materials 12(3):300–313. https://doi.org/10.1177/002199837801200307

[48] Shrivastava S, Mohite PM, Yadav T, et al (2018) Multi-objective multi-laminate design and optimization of a carbon fibre composite wing torsion box using evolutionary algorithm. Compos Struct 185:132–147. https://doi.org/10.1016/j.compstruct.2017.10.041

[49] Sørensen R, Lund E (2015) Thickness filters for gradient based multi-material and thickness optimization of laminated composite structures. Struct Multidiscip Optim 52:227–250. https://doi.org/10.1007/s00158-015-1230-3

[50] Sorour SS, Saleh CA, Shazly M (2024) A review on machine learning implementation for predicting and optimizing the mechanical behaviour of laminated fiber-reinforced polymer composites. Heliyon 10:e33681. https://doi.org/10.1016/j.heliyon.2024.e33681

[51] Terada Y, Todoroki A, Shimamura Y (2001) Stacking sequence optimizations using fractal branch and bound method for laminated composites. JSME Int J Ser A 44(4):490–498. https://doi.org/10.1299/jsmea.44.490

[52] Tsai SW, Wu EM (1971) A general theory of strength for anisotropic materials. Journal of Composite Materials 5(1):58–80. https://doi.org/10.1177/002199837100500106

[53] Wang X, Jin Y, Schmitt S, et al (2023) Recent advances in Bayesian optimization. ACM Computing Surveys 55(13s):1–36. https://doi.org/10.1145/3582078