

Zero-shot rib design: merging a training-free generative prior with topology optimization

Yongmin Kwon^{a,b}, Namwoo Kang^{a,b,*}

^aCho Chun Shik Graduate School of Mobility, Korea Advanced Institute of Science and Technology, 193 Munji-ro, Yuseong-gu, Daejeon, 34051, Republic of Korea

^bAI Research Team, Narnia Labs, 193 Munji-ro, Yuseong-gu, Daejeon, 34051, Republic of Korea

ARTICLE INFO

Keywords:

Topology optimization
Score distillation sampling
Text-guided structural design
Rib reinforcement design
Training-free generative design
Design intent representation

ABSTRACT

Natural load-bearing patterns such as leaf venation, trabecular bone, and spider webs achieve high stiffness per unit mass, yet classical topology optimizers rarely reach such geometries, and few let engineers express structural design intent through natural language. This work treats a frozen text-to-image diffusion model as a training-free source of design knowledge and distills it into the physics loop of density-based topology optimization via score distillation sampling, so that a text prompt becomes an explicit, machine-interpretable representation of engineer intent. The prompt-induced generative gradient and the finite element sensitivity are combined at every iteration, letting physics decide which prompt-induced features survive. In 245 primary SDS runs spanning four geometric domains and two physics regimes, 38 of 49 prompt–domain combinations achieved statistically significant compliance reductions (up to -31.5% mechanical and -23.0% thermoelastic), outperforming gradient-based baselines. Cross-domain morphological analysis identifies a recurring structural signature of improvement: in most domains the generative prior suppresses dead-end branches in the rib skeleton, with endpoint–compliance correlation $r = +0.56$ to $+0.99$. A Heaviside projection with β -continuation resolves a pronounced intermediate-density tendency in this diffusion–physics coupling (42.6% to $< 3\%$), and an automated skeleton-based pipeline converts optimized density fields into candidate geometry ready for computer-aided design. By retargeting the generative prior across domains, loading conditions, and physics objectives through a change of text prompt, with each new problem’s physics setup specified separately, the framework uses a pretrained generative model as a reusable, training-free prior for engineering design.

1. Introduction

Topology optimization determines the optimal distribution of material within a prescribed design domain to maximize structural performance subject to physical and manufacturing constraints [1]. The Solid Isotropic Material with Penalization (SIMP) method is the most widely adopted framework. A continuous density variable $\rho_e \in [0, 1]$ is assigned to each finite element and updated by gradient-based optimization to minimize structural compliance [2–5]. Recent isogeometric formulations employ multi-patch, T-spline, and Bézier-extraction representations [6–8] to handle complex plate and shell geometries with smoother boundaries, while particle-flow and material-point formulations [9, 10] provide alternative Lagrangian descriptions well suited to large geometric changes and hinge-free compliant-mechanism design. Despite its maturity, SIMP suffers from a fundamental limitation: the non-convexity introduced by the penalization parameter generates a rugged objective landscape with numerous local minima, so gradient-based optimizers converge to different solutions depending on the initial density distribution, with no guarantee of approaching the global optimum [11, 12].

This limitation is particularly acute in the design of rib reinforcements for thin-walled plate structures. Rib stiffeners serve as key elements for structural efficiency in virtually all thin-walled artifacts, including automotive tailgate panels, aircraft fuselage and wing skins, electronic device housings, and ship decks [13–15]. A rib is a thin wall protruding from one side of a plate, which substantially enhances local bending

*Corresponding author.

✉ kymin1002@kaist.ac.kr (Y. Kwon); nwkang@kaist.ac.kr (N. Kang)

stiffness while minimizing weight increase. However, determining the optimal rib layout is a combinatorial, knowledge-intensive design task: the number, orientation, height, spacing, and branching topology of ribs all constitute design variables, producing an especially challenging optimization landscape [16]. Nature offers many efficient rib-like patterns (for example, leaf venation, trabecular bone, and dragonfly wing venation), many of which may not lie near the local optima identified by SIMP and are therefore unlikely to be reached by its local search. Multi-start strategies that launch optimization from multiple random initializations can mitigate the problem, but the exploration remains unstructured and the computational cost scales linearly. A method that escapes local optima while preserving physical optimality is therefore needed.

What conventional methods lack is a way to make design knowledge explicit and machine-interpretable, so that an optimizer can act on it directly. Expert designers draw on many structural patterns for rib layout, such as radial webs, branching venation, and cellular tessellations from other disciplines and from nature, yet these patterns have lacked a representation the optimizer can consume. The rib networks sought here are themselves nature-inspired, approximating the load-bearing geometries seen in leaf venation and trabecular bone. A diffusion model trained on billions of natural images therefore already encodes much of the structural vocabulary that rib design draws on, which makes it a well-matched, and so far unused, source of design knowledge for this task. Communicating design intent through generated text has long been studied in collaborative engineering design [17], and large language models are rapidly expanding the role of natural language in mechanics, product design, and manufacturing [18]. What has been missing is a mechanism that turns such linguistic intent into physics-coupled optimization gradients. Parametric templates capture only a narrow, pre-enumerated slice of this knowledge, and data-driven generative models internalize it implicitly in network weights tied to a single training distribution (Section 2). This work instead represents structural design intent as a natural-language prompt. The visual prior of a frozen text-to-image diffusion model [19] is distilled through score distillation sampling (SDS) [20] and coupled to the finite element sensitivity of SIMP at the gradient level. The prompt thereby becomes a direct input to the optimization: an explicit knowledge representation whose physical consequences are filtered, at every iteration, by finite element analysis (FEA). No prior study has applied SDS to density-based topology optimization with generative priors coupled to FEA at the gradient level.

The framework proposed here addresses this gap through three contributions.

First, this constitutes the first application of score distillation sampling to physics-constrained topology optimization. By coupling a frozen Stable Diffusion 2.1 [19] to SIMP at the gradient level, the framework enables text-guided structural design without any training data or fine-tuning, a capability absent from existing data-driven generative methods. The framework is training-free. It needs no task-specific training, fine-tuning, or labeled topology-optimization data. Instead, it builds on a large-scale, publicly pretrained diffusion model whose weights stay frozen throughout optimization. Retargeting the generative prior to a new domain, loading condition, or physics objective needs no model retraining, though each new engineering problem still requires its own geometry, boundary and loading conditions, physical model, constraints, and numerical parameters. This extends the training-free paradigm of DreamFusion [20] from vision to structural engineering and turns a pretrained foundation model into a reusable knowledge source for engineering design. The prompt is the only change to the generative prior. As in any topology-optimization study, each new problem still requires the design domain, boundary and loading conditions, physical model, volume fraction, and numerical parameters, including the SDS weight λ_{SDS}^0 (Section 3.3). The retargeting cost is therefore much lower than for data-driven models that require retraining.

Second, dead-end suppression is identified as the predominant structural signature of SDS-guided improvement. Cross-domain morphological analysis across six physics settings reveals a positive correlation between skeleton endpoint count and compliance (Pearson $r = +0.56$ to $+0.99$, median $r = +0.97$), indicating that the generative prior improves designs through one physically interpretable mechanism. This cross-domain finding clarifies the mechanism: in most domains the gain comes from removing non-load-bearing dead ends, not from adding branches.

Third, the framework diagnoses and resolves a pronounced gray-density tendency in the SDS-SIMP coupling and is validated for generality and scalability. Heaviside projection with β -continuation is itself a standard tool in density-based optimization. The contribution here is not the projection but the diagnosis that SDS induces a pronounced gray-density tendency (roughly doubling

the intermediate-density ratio in the present SDS-SIMP coupling), resolved by synchronizing the β -schedule with the SDS schedule. This reduces intermediate densities from 42.6% to below 3%, and an automated skeleton-based pipeline converts the optimized density field directly into computer-aided design (CAD) geometry, i.e., a CAD-convertible candidate. Validation spans 245 primary SDS runs across four geometric domains (synthetic benchmarks through an industrial automotive suspension link) and two physics regimes (mechanical and thermoelastic bending). Of 49 primary prompt-domain combinations, 38 achieved statistically significant compliance reductions ($p = 0.031$ for the mechanical subset), with gains up to -31.5% mechanical and -23.0% thermoelastic that consistently outperform all alternative baselines. A negative result on heat conduction (Section 7) delineates the boundary of applicability.

The remainder of this paper is organized as follows. Section 2 reviews related work on data-driven generative topology optimization and score-based priors. Section 3 details the proposed framework, and Section 4 the experimental setup. Section 5 reports results across the four domains and two physics regimes, and Section 6 analyzes the mechanisms underlying the observed improvements. Section 7 discusses implications and limitations, and Section 8 concludes.

2. Related work

2.1. Data-driven generative models for topology optimization

Recent work has pursued data-driven generative models to overcome the local-optima limitation of classical topology optimization. TopoDiff [21] trained a diffusion model on approximately 33,000 topology-optimization solutions, using von Mises stress, strain energy density, and volume fraction as conditioning channels, achieving an eightfold reduction in compliance error compared to earlier generative adversarial network-based approaches. Diffusion Optimization Models (DOM) [22] learn entire SIMP optimization trajectories rather than single optimal designs, exploiting the structural similarity between iterative denoising and iterative convergence. The Physics-Informed Diffusion Model (PIDM) [23] integrates partial differential equation (PDE) residuals into training to reduce physical constraint violations by two orders of magnitude. Extensions to three-dimensional (3D) problems [24, 25], to large-scale structural datasets [26, 27], and to manufacturability-aware deep generative design [28] have also been explored. Learning-based components have also been integrated into topology optimization, including differentiable multiscale design [29] and network-based robust or concurrent formulations [30–32]. These still rely on task-specific training or surrogate data tied to a particular design distribution, unlike the frozen, publicly pretrained prior coupled to FEA at the gradient level in the present work. Recent latent-diffusion pipelines [33, 34] still learn their priors from precomputed solutions and thus inherit the dataset dependence discussed below, whereas training-dataset-free reparameterization [35] avoids the dataset but injects no external design knowledge or language interface.

The representative data-driven methods considered here retain, to varying degrees, three limitations. First, they require massive domain-specific training datasets (tens of thousands of samples) tied to specific load and boundary conditions. Data-efficiency remedies such as geometric augmentation guided by engineering uncertainty [36] alleviate but do not remove this dependence. Second, they lose generalization when presented with unfamiliar conditions outside the training distribution. Third, the generative model produces outputs independently of physical analysis, offering no intrinsic guarantee of compliance optimality. These limitations apply unevenly, and several recent methods partially address them. PIDM [23] embeds PDE residuals during training to improve physical consistency, and DOM [22] learns full optimization trajectories rather than static optima. Prior methods, however, incorporate physics primarily at training time. Among the methods reviewed in Tables 1 and 2, none appears to couple a frozen text-to-image score gradient directly with an FEA sensitivity at every density-update iteration, so the evolving design is not tested against physical optimality during generation. The present framework closes this gap. Table 1 contrasts the training and retargeting costs of these methods with the present framework.

2.2. Score distillation and text-driven structural design

In a parallel development in computer vision, score distillation sampling, introduced by DreamFusion [20], demonstrated that the visual prior of a frozen text-to-image model (e.g., Stable Diffusion [19]) can be distilled into a 3D optimization loop, a Neural Radiance Field (NeRF) [37], without any training data,

Table 1

Comparison of data-driven diffusion-based topology optimization (TO) methods.

	TopoDiff [21]	DOM [22]	PIDM [23]	This work
Training data	~33,000 TO solutions	~10,000 trajectories	PDE-specific	0 (zero-shot)
Est. training cost	~2,750 GPU-hrs	~830 GPU-hrs	Physics training	0
New domain	Full rebuild	Full rebuild	Retraining required	No retraining
New load condition	Full rebuild	Full rebuild	Retraining required	No retraining

achieving high-quality text-to-3D generation. The core principle is that the score function of the pretrained diffusion model yields pixel-level gradients that indicate how the current output should be modified to better align with the text prompt. SDS-type score priors have also been applied to inverse problems such as medical imaging [38–40], similarly leveraging diffusion knowledge to structure an optimization problem. Separately, Contrastive Language-Image Pre-training (CLIP)-based topology optimization [41] used contrastive embeddings to impose text-based styles on optimized structures, but CLIP provides only a global image–text similarity measure and cannot generate spatially detailed generative gradients at the pixel level. In product design more broadly, knowledge-graph-guided semantic diffusion has been used for bio-inspired form generation [42], though without coupling to physical analysis. Closest in spirit to the present work, the recently proposed LMTO framework [43] couples a large visual–language model (LVLM) to conceptual topology optimization, semantically decomposing candidate structures within the large model’s knowledge space and steering the optimization toward human preference. LMTO and the present framework are complementary: LMTO uses the large model interpretively at the semantic level, whereas the present framework distills the generative prior into pixel-level score gradients summed with the FEA sensitivity at every density update. The two could in principle be combined. Table 2 contrasts all text- and diffusion-based approaches.

Table 2

Comparison of diffusion- and text-based approaches for structural design.

	Data-driven diffusion TO [21, 22]	SDS for vision [20]	CLIP-based TO [41]	LVLM-based TO [43]	This work
Training data	Tens of thousands of TO solutions	Not required	Not required	Not required	Not required
Physics integration	Condition channel (passive)	None	Compliance (objective)	Preference-weighted loop	FEA sensitivity (active grad.)
Design guidance	Conditional generation	Text → 3D	Global similarity (CLIP)	Semantic decomposition & preference weighting	Pixel-level score gradient
Output / domain	Pixel density (TO)	NeRF (3D vision)	Density field (TO style)	Density field (conceptual TO)	Element density (rib design)
New-condition adaptability	Full retraining	High (prompt)	High (prompt)	High (concept)	High (no retraining)

2.3. Research gap

Domain-specific generative models remain valuable when abundant labeled data exist, but their applicability is tied to the training distribution: a model trained under specific boundary conditions cannot transfer to other physics problems or unfamiliar domain geometries without retraining. Worse, because generated designs are produced at generation time without active coupling to an FEA solver, nothing guarantees they satisfy physical equilibrium, meet volume constraints, or approach optimality by any quantitative criterion. Even methods that embed physics during training may yield outputs that are visually plausible yet physically suboptimal. From a knowledge-representation standpoint, the design knowledge such

models encode is implicit, non-transferable, and inaccessible to the engineer at design time. In summary, existing approaches fall into three groups: those that embed design knowledge implicitly in retraining-bound network weights, those that impose text guidance without spatially detailed physics-coupled gradients, and those that use a large model interpretively to weight the optimization toward semantic preference. Gradient-level coupling between a frozen text-to-image prior and an FEA sensitivity remains unexplored for density-based topology optimization. Closing this gap is the focus of the present work.

3. Methodology

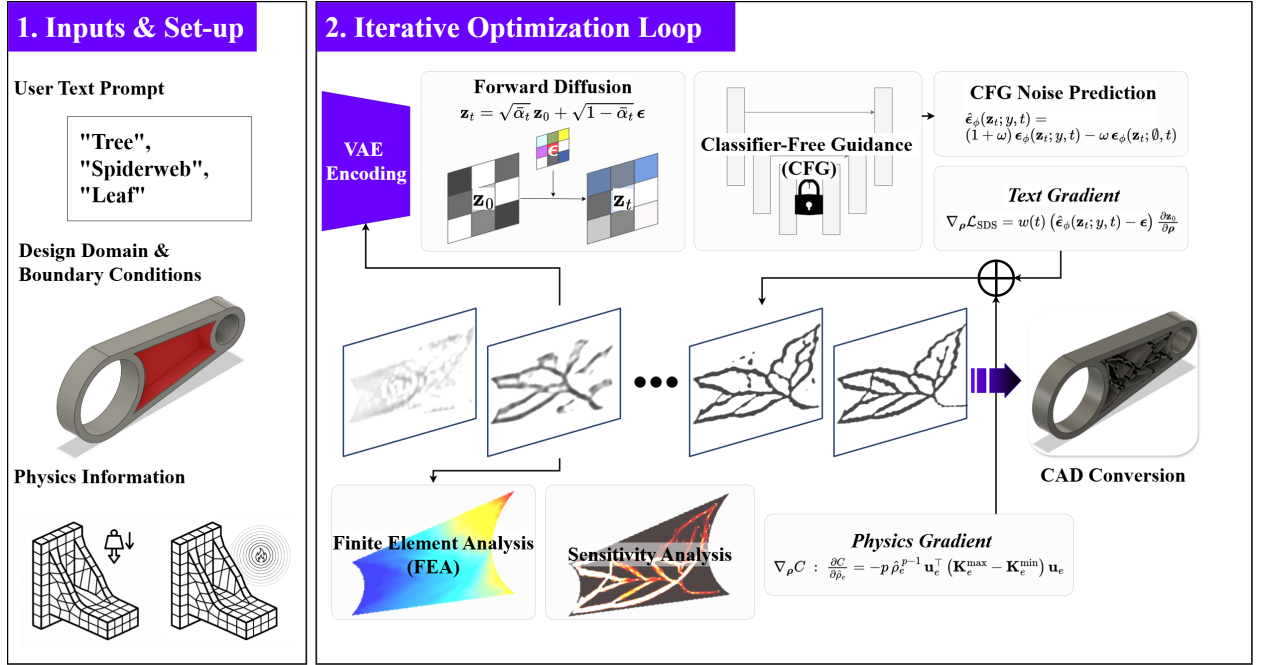


Figure 1: Computational framework for text-guided, physics-constrained topology optimization. Three stages: **(i) Inputs & Set-up** (text prompt, design domain, boundary conditions, volume fraction; frozen Stable Diffusion 2.1 as generative prior); **(ii) Iterative Optimization Loop** coupling a Mindlin–Reissner plate FEA physics gradient with an SDS text gradient under four-stage scheduling; **(iii) CAD Conversion** of the converged density field into CAD-convertible candidate geometry.

The framework integrates two gradient sources into a single optimization loop that updates the density field $\hat{\rho}$ (Fig. 1). The **Inputs & Set-up** phase specifies a user text prompt, the design domain with boundary conditions, and the physics information. At each iteration of the **Iterative Optimization Loop**, a Mindlin–Reissner plate FEA solver produces the physics gradient $\nabla_\rho C$. In parallel, the current density field is encoded into the latent space of a frozen Stable Diffusion 2.1, forward diffusion adds noise ϵ , and the U-Net predicts this noise under classifier-free guidance (CFG). The discrepancy between predicted and actual noise yields the SDS text gradient $\nabla_\rho \mathcal{L}_{\text{SDS}}$ via backpropagation. The two gradients combine as a weighted sum, Eq. (6), under a four-stage schedule governing their interaction:

- **Warmup** ($k_{\text{warm}} = 30$): only the physics gradient acts, stabilizing initial load paths.
- **Cosine cooldown** ($k_{\text{cool}} = 500$): the SDS weight $\lambda_{\text{SDS}}(k)$ decays gradually to zero.
- **Timestep annealing**: the diffusion timestep t linearly decreases from $t_{\text{max}} = 0.98$ to $t_{\text{min}} = 0.50$ for coarse-to-fine structural guidance.
- **Exponential moving average (EMA) smoothing**: an EMA of the SDS gradient damps single-sample stochastic noise.

The *Adam* optimizer updates the density field under a bisection-based volume constraint and Heaviside projection with β -continuation for binarization. In the **CAD Conversion** stage, the converged density field is automatically processed by binarization, medial-axis thinning, Bézier curve fitting, and extrusion to yield CAD-convertible candidate geometries.

3.1. Problem formulation

The design problem is formulated as a two-dimensional (2D) density-based topology optimization. The design domain Ω is discretized into an $n_{\text{elx}} \times n_{\text{ely}}$ finite element mesh, with a continuous density variable $\rho_e \in [0, 1]$ assigned to each element. The optimization problem minimizes structural compliance subject to a volume fraction constraint:

$$\begin{aligned} \min_{\boldsymbol{\rho}} \quad & C(\boldsymbol{\rho}) = \mathbf{F}^\top \mathbf{U} = \sum_{e=1}^N (\hat{\rho}_e)^p \mathbf{u}_e^\top \mathbf{k}_e \mathbf{u}_e \\ \text{s.t.} \quad & \mathbf{K}(\hat{\boldsymbol{\rho}}) \mathbf{U} = \mathbf{F}, \\ & \frac{V(\hat{\boldsymbol{\rho}})}{|\Omega|} = V^*, \\ & 0 \leq \rho_e \leq 1 \quad \forall e \end{aligned} \tag{1}$$

where C is the compliance, $\mathbf{K}$, $\mathbf{U}$, and $\mathbf{F}$ are the global stiffness matrix, nodal displacement vector, and external force vector, respectively, and V^* is the target volume fraction. The density field passes through three stages. The *raw design variables* $\boldsymbol{\rho}$ are updated directly by the optimizer. The *filtered density* $\tilde{\boldsymbol{\rho}} = \mathcal{F}(\boldsymbol{\rho})$ is obtained by a spatial averaging density filter [44, 45] with radius $r_{\min} = 5$ that removes checkerboard patterns. Because this radius is fixed in physical units, the filter also limits mesh-dependent features (Appendix D.7). The *physical density* $\hat{\boldsymbol{\rho}} = \mathcal{H}(\tilde{\boldsymbol{\rho}})$ is produced by a smoothed Heaviside projection [46] that drives elements toward near-binary $\{0, 1\}$ material states. The sequence $\boldsymbol{\rho} \xrightarrow{\mathcal{F}} \tilde{\boldsymbol{\rho}} \xrightarrow{\mathcal{H}} \hat{\boldsymbol{\rho}}$ is fully differentiable. Compliance, stiffness, and all sensitivities are evaluated on the physical density $\hat{\boldsymbol{\rho}}$, and chain-rule backpropagation carries the gradient back to $\boldsymbol{\rho}$. Under SIMP interpolation [2], the element stiffness is computed as $\mathbf{K}_e(\hat{\rho}_e) = \hat{\rho}_e^p \mathbf{K}_e^{\max} + (1 - \hat{\rho}_e^p) \mathbf{K}_e^{\min}$ with penalty $p = 3$, where $\mathbf{K}_e^{\max}$ and $\mathbf{K}_e^{\min}$ correspond to elements with and without rib reinforcement, respectively. Since base-plate regions retain physical substance ($\mathbf{K}_e^{\min} > 0$), the formulation avoids the unphysical assumption that zero-density regions bear no load. Detailed derivations of the SIMP interpolation, density filter, and Heaviside projection are provided in Appendices A.2, A.3, and A.5.

3.2. Finite element analysis

The three plate-bending domains (D_{Rect} , D_{Circle} , D_{Hole}) are analyzed using Mindlin–Reissner plate theory [47, 48] with the Mixed Interpolation of Tensorial Components (MITC4) element [49]. The Mindlin–Reissner theory is adopted because rib-reinforced plates behave locally as thick plates in the rib regions, where transverse shear deformation ($\gamma \neq 0$) cannot be neglected as in the Kirchhoff thin-plate theory. The MITC4 element is a 4-node bilinear quadrilateral with three degrees of freedom (DOF) per node (w, θ_x, θ_y), whose element stiffness matrix comprises bending and shear contributions:

$$\mathbf{K}_e = \int_{\Omega_e} \mathbf{B}_b^\top \mathbf{D}_b \mathbf{B}_b d\Omega + \int_{\Omega_e} \mathbf{B}_s^{\text{MITC}\top} \mathbf{D}_s \mathbf{B}_s^{\text{MITC}} d\Omega \tag{2}$$

where $\mathbf{D}_b$ is the bending constitutive matrix and $\mathbf{D}_s$ the shear constitutive matrix (with Reissner correction factor $\kappa_s = 5/6$). The MITC4 formulation overcomes shear locking by evaluating shear strains at four tying points on the element edges and interpolating them into the interior, yielding accurate solutions even in the thin-plate limit.

In the rib-reinforced plate model, zero-density regions retain the physical substance of the base plate rather than being void. A modified SIMP interpolation directly interpolates bending and shear stiffnesses: regions with ribs have effective thickness $t_{\max}$ (base plate + rib height), while regions without ribs have only $t_{\min}$ (base plate). Since flexural rigidity scales as t^3 and shear stiffness as t , the element stiffness is interpolated as $\mathbf{K}_e(\hat{\rho}_e) = \hat{\rho}_e^p \mathbf{K}_e^{\max} + (1 - \hat{\rho}_e^p) \mathbf{K}_e^{\min}$, where $\mathbf{K}_e^{\max}$ and $\mathbf{K}_e^{\min}$ can be precomputed. The industrial D_{Link}

domain is analyzed under plane stress using a standard 4-node bilinear quadrilateral element with two degrees of freedom per node (u, v).

In both formulations, the compliance sensitivity exploits self-adjointness: $\partial C / \partial \hat{\rho}_e = -p \hat{\rho}_e^{p-1} \mathbf{u}_e^\top (\mathbf{K}_e^{\max} - \mathbf{K}_e^{\min}) \mathbf{u}_e$, where the quantity $\mathbf{u}_e^\top (\mathbf{K}_e^{\max} - \mathbf{K}_e^{\min}) \mathbf{u}_e$ represents the strain energy associated with the stiffness gain from adding a rib to element e . Elements with larger values yield greater compliance reduction upon rib placement. The sensitivity is extended to the raw design variables via the chain rule through the Heaviside projection and density filter. Checkerboard patterns and mesh dependency are suppressed by the density filter ($r_{\min} = 5$), which spatially averages raw densities within a neighborhood and provides indirect control of the minimum feature size. A companion maximum-thickness filter ($r_{\max} = 10$) additionally caps local rib thickness for manufacturability through a quadratic local-volume penalty added to the physics update. Its formulation and sensitivity are given in Appendix A.4.

For the thermoelastic experiments, the same MITC4 plate FEA is used with a different load vector: a prescribed uniform through-thickness temperature gradient ΔT induces thermal bending moments $\mathbf{M}_{\text{th}} = \alpha_T \Delta T [1, 1, 0]^\top$, with thermal expansion coefficient α_T , on every element, which are converted to equivalent nodal forces $\mathbf{F}_{\text{th}}$ via Gauss integration. In contrast to the design-dependent thermal loads of material-based thermoelastic formulations [50], a simplified prescribed-thermal-moment weak coupling is adopted. The thermal moment $\mathbf{M}_{\text{th}}$ is set by the prescribed ΔT alone and does not scale with the element stiffness $E(\hat{\rho})$. The assembled load $\mathbf{F}_{\text{th}}$ is therefore density-independent ($\partial \mathbf{F}_{\text{th}} / \partial \hat{\rho} = \mathbf{0}$), so the compliance sensitivity keeps its standard self-adjoint form, $\partial C / \partial \hat{\rho}_e = -\mathbf{u}_e^\top (\partial \mathbf{K}_e / \partial \hat{\rho}_e) \mathbf{u}_e$. A finite-difference check is provided in Appendix A.7. Full derivations of the element formulations, constitutive matrices, and sensitivity are given in Appendices A.1, A.2, and A.6.

3.3. Score distillation sampling for topology optimization

The physical density field $\hat{\rho} \in \mathbb{R}^{n_{\text{elx}} \times n_{\text{ely}}}$ is converted to a three-channel RGB image by replicating the single grayscale channel, resized to 512×512 via bilinear interpolation, and mapped to the latent space by the encoder $\mathcal{E}$ of the frozen variational autoencoder (VAE) of Stable Diffusion 2.1, giving the latent code $\mathbf{z}_0 = \mathcal{E}(\mathbf{x}) \in \mathbb{R}^{4 \times 64 \times 64}$. This entire pipeline is differentiable, enabling backpropagation to $\hat{\rho}$ (Appendix B.1). Fig. 2 illustrates the full SDS pipeline.

Forward diffusion adds Gaussian noise to the latent vector at a randomly sampled timestep t :

$$\mathbf{z}_t = \sqrt{\bar{\alpha}_t} \mathbf{z}_0 + \sqrt{1 - \bar{\alpha}_t} \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \quad (3)$$

where $\bar{\alpha}_t = \prod_{s=1}^t (1 - \beta_s)$ is the cumulative noise schedule [51]. The U-Net noise predictions for the text prompt y and the null condition are combined via CFG [52]:

$$\hat{\boldsymbol{\epsilon}}_\phi(\mathbf{z}_t, y, t) = (1 + \omega) \boldsymbol{\epsilon}_\phi(\mathbf{z}_t, y, t) - \omega \boldsymbol{\epsilon}_\phi(\mathbf{z}_t, \emptyset, t) \quad (4)$$

with guidance scale $\omega = 100$. The SDS gradient is computed with a stop-gradient on the U-Net Jacobian:

$$\nabla_{\boldsymbol{\rho}} \mathcal{L}_{\text{SDS}} = w(t) (\hat{\boldsymbol{\epsilon}}_\phi(\mathbf{z}_t, y, t) - \boldsymbol{\epsilon}) \frac{\partial \mathbf{z}_0}{\partial \boldsymbol{\rho}} \quad (5)$$

where $w(t) = 1 - \bar{\alpha}_t$ is a time-dependent weight. The physics and SDS gradients are combined as a weighted sum:

$$\nabla_{\boldsymbol{\rho}} \mathcal{L}_{\text{total}} = \lambda_{\text{phys}} \nabla_{\boldsymbol{\rho}} C + \lambda_{\text{sds}}(k) \nabla_{\boldsymbol{\rho}} \mathcal{L}_{\text{SDS}} \quad (6)$$

For clarity, Eq. (6) shows only the two principal gradients. The physics term additionally carries the maximum-thickness penalty gradient $\nabla_{\boldsymbol{\rho}} \mathcal{L}_{\text{max}}$ used for manufacturability (Appendix A.4). The weights are $\lambda_{\text{phys}} = 1.0$, while λ_{sds}^0 is scaled to the magnitude of the physics loss so that the generative signal stays competitive in the raw weighted sum. This sets $\lambda_{\text{sds}}^0 = 10$ for the mechanical point-load domains D_{Rect} , D_{Circle} , and D_{Link} (baseline compliance ~ 3 –9), $\lambda_{\text{sds}}^0 = 50$ for D_{Hole} (baseline compliance ~ 63 , roughly seven times higher due to 4-sided loading around the central hole), and $\lambda_{\text{sds}}^0 = 100$ for thermoelastic compliance (baseline compliance ~ 600 , with distributed thermal moments on every element. Empirical sweep in Appendix D.6 and Table 10). In each case, raising λ_{sds}^0 proportionally to the physics-loss scale

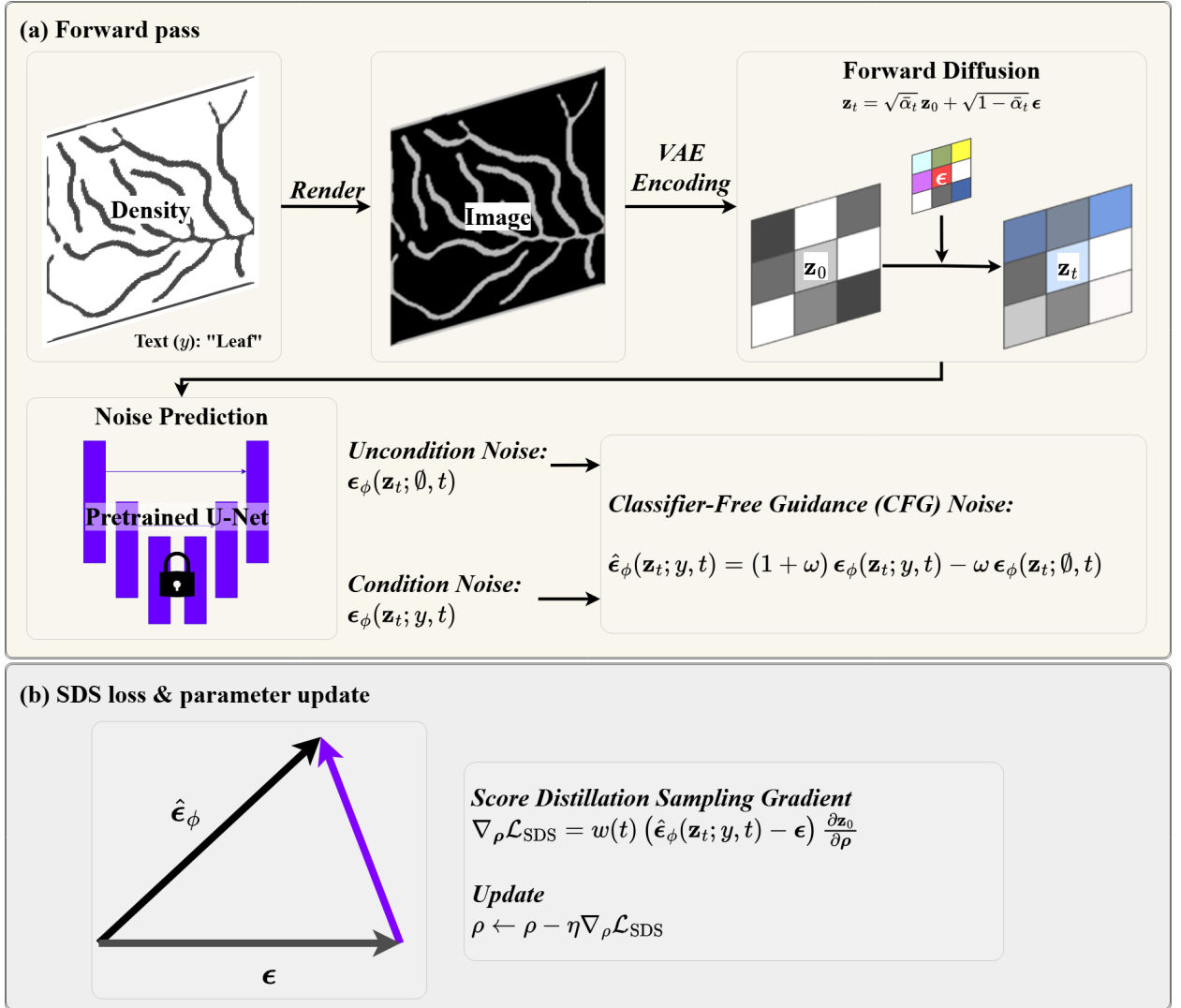


Figure 2: SDS pipeline for topology optimization. (a) Forward pass: the density field $\hat{\rho}$ is rendered as a grayscale image and encoded into the latent $\mathbf{z}_0$ by the frozen VAE; Gaussian noise ϵ is added at a sampled diffusion timestep t to produce $\mathbf{z}_t$; the frozen U-Net predicts the noise under both the text-conditioned input (prompt y) and the unconditional (null) input; the two predictions are combined via CFG with weight ω to form $\hat{\epsilon}_\phi$. (b) SDS loss and parameter update: the difference $\hat{\epsilon}_\phi - \epsilon$ is weighted by $w(t) = 1 - \bar{\alpha}_t$ and backpropagated through the differentiable encode-render chain $\partial \mathbf{z}_0 / \partial \rho$ to yield the SDS gradient $\nabla_\rho \mathcal{L}_{\text{SDS}}$, which is combined with the physics gradient in the final density update.

preserves a comparable relative influence between the physics and SDS gradients in the raw-sum update. Eq. (6) is a heuristic gradient-blending update rather than the exact gradient of a single scalar objective, with $\lambda_{\text{sds}}(k)$ setting the trade-off between the compliance and generative terms [20] (the SDS term uses a stop-gradient approximation and is not the exact gradient of a scalar loss). Its behaviour is controlled empirically through gradient-scale matching, a decaying SDS weight, exact volume projection, and a final physics-only convergence phase: the cosine cooldown drives $\lambda_{\text{sds}}(k) \rightarrow 0$, so the final iterations follow the pure physics gradient toward a physically consistent optimum.

3.4. Scheduling strategy

Naively applying the SDS gradient at constant intensity throughout optimization causes two problems: (1) before the physics solver establishes load paths in the early stages, SDS distorts the density field and generates

unphysical patterns, and (2) in the final stages, stochastic SDS noise interferes with fine convergence. A four-stage scheduling strategy is therefore designed.

During the warmup phase (iterations 1 to $k_{\text{warm}} = 30$), the SDS gradient is deactivated and only the physics gradient is used, forming a basic material distribution between external forces and supports that SDS subsequently refines. Since the Heaviside projection is weak at $\beta = 1$ during the first 50 iterations, the density field evolves smoothly and exploratorily during this period.

After warmup, the SDS weight follows a cosine cooldown:

$$\lambda_{\text{SDS}}(k) = \lambda_{\text{SDS}}^0 \cdot \frac{1}{2} \left(1 + \cos \left(\frac{\pi \cdot \min(k - k_{\text{warm}}, k_{\text{cool}})}{k_{\text{cool}}} \right) \right) \quad (7)$$

with $k_{\text{cool}} = 500$, so that the SDS influence converges to zero after approximately 530 iterations and the final convergence proceeds with the physics gradient alone. Cosine decay maintains the SDS influence longer in the early stages compared to linear decay while decreasing sharply in the later stages, enabling a natural transition from exploration (SDS-guided) to exploitation (physics-driven).

Simultaneously, the diffusion timestep is linearly annealed:

$$t(k) = t_{\text{max}} - (t_{\text{max}} - t_{\text{min}}) \cdot \frac{k - k_{\text{warm}}}{k_{\text{cool}}} \quad (8)$$

from $t_{\text{max}} = 0.98$ to $t_{\text{min}} = 0.50$, implementing a coarse-to-fine strategy [53]. Noise predictions at high t induce global geometry (the coarse arrangement of the overall pattern, symmetry, and connectivity), while predictions at low t refine local details such as branching points, member thickness, and boundary smoothness.

Finally, the SDS gradient is stabilized by EMA smoothing:

$$\mathbf{g}_{\text{ema}}^{(k)} = \alpha \mathbf{g}_{\text{ema}}^{(k-1)} + (1 - \alpha) \nabla \mathcal{L}_{\text{SDS}}^{(k)} \quad (9)$$

with $\alpha = 0.9$, averaging gradients over approximately the most recent 10 iterations to damp the stochastic variance inherent in single-sample SDS gradient estimation (Appendix B.2).

3.5. Optimization loop and implementation

The *Adam* optimizer [54] is used with momentum coefficients $\beta_1 = 0.5$ and $\beta_2 = 0.9$ and learning rate annealed from 0.04 to 0.01. *Adam* is preferred over the conventional Optimality Criteria (*OC*) update for two reasons: it accommodates arbitrary gradient combinations (physics + SDS), and it converges stably in conjunction with the Heaviside projection, whereas *OC* exhibits period-2 oscillation at high β that prevents convergence (Appendix B.4; quantitative baseline comparison in Section 5.2). A bisection projection enforces the volume fraction constraint exactly at each iteration (Appendix B.3). Algorithm 1 summarizes the complete optimization loop, and Table 3 collects the hyperparameter values used throughout the study.

3.6. Post-processing

The optimized density field is automatically converted into CAD-convertible candidate geometry in the Standard for the Exchange of Product model data (STEP) format through a six-stage pipeline. The continuous density field is binarized at threshold $\rho_{\text{th}} = 0.5$. Zhang–Suen thinning [55] then extracts the medial-axis skeleton, and a Euclidean distance transform computes local rib half-widths. The skeleton is converted to a topological graph with branch-point and endpoint nodes, each edge is fitted with piecewise cubic Bézier curves by least squares, and offset polygons of physically meaningful width are generated from the fitted centerlines. Finally, the 2D polygons are extruded into 3D solids and exported as AP214-schema STEP files. The entire pipeline completes in 5 seconds on a single CPU core and requires no user parameters beyond the extrusion height. The exported STEP solids are watertight and importable into standard CAD software as “CAD-convertible candidates,” their fabrication feasibility assessed downstream rather than verified here. A 3D solid re-analysis (Appendix D.8) confirms the 2D design rankings are largely preserved. Full pipeline details are provided in Appendix C.

Algorithm 1 Zero-Shot Rib Design

Require: Domain Ω , boundary conditions, V^* , text prompt y
Ensure: Optimized density field $\hat{\rho}$

- 1: Initialize $\rho_e = V^*$ for all $e \in \Omega_{\text{active}}$ {Uniform density}
- 2: Load pretrained Stable Diffusion 2.1 (all weights frozen)
- 3: $\beta \leftarrow 1.0$ {Initial Heaviside sharpness}
- 4: $\mathbf{g}_{\text{ema}} \leftarrow \mathbf{0}$ {EMA buffer}
- 5: **for** $k = 1$ to $K = 1000$ **do**
- 6: $\tilde{\rho} \leftarrow \text{DensityFilter}(\rho, r_{\min})$
- 7: $\hat{\rho} \leftarrow \text{HeavisideProject}(\tilde{\rho}, \beta)$
- 8: Solve $\mathbf{K}(\hat{\rho})\mathbf{u} = \mathbf{f}$ via sparse Cholesky {FEA}
- 9: $\mathbf{g}_{\text{phys}} \leftarrow \partial C / \partial \rho + \partial \mathcal{L}_{\max} / \partial \rho$ {compliance + max-thickness penalty (App. A.4)}
- 10: **if** $k > k_{\text{warm}}$ **and** $\lambda_{\text{sds}}(k) > 0$ **then**
- 11: $\mathbf{x} \leftarrow \text{resize}(\text{stack_rgb}(\hat{\rho}), 512 \times 512)$
- 12: $\mathbf{z}_0 \leftarrow \mathcal{E}(\mathbf{x})$ {VAE encoding}
- 13: $t \leftarrow \text{anneal}(k, t_{\max}, t_{\min})$
- 14: $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
- 15: $\mathbf{z}_t \leftarrow \sqrt{\alpha_t} \mathbf{z}_0 + \sqrt{1 - \alpha_t} \epsilon$ {Forward diffusion}
- 16: $\hat{\epsilon} \leftarrow (1 + \omega) \epsilon_{\phi}(\mathbf{z}_t, y, t) - \omega \cdot \epsilon_{\phi}(\mathbf{z}_t, \emptyset, t)$ {CFG}
- 17: $\mathbf{g}_{\text{sds}}^{\text{raw}} \leftarrow w(t)(\hat{\epsilon} - \epsilon) \partial \mathbf{z}_0 / \partial \rho$ {SDS gradient}
- 18: $\mathbf{g}_{\text{ema}} \leftarrow \alpha \cdot \mathbf{g}_{\text{ema}} + (1 - \alpha) \cdot \mathbf{g}_{\text{sds}}^{\text{raw}}$
- 19: $\lambda_{\text{sds}} \leftarrow \text{CosineDecay}(k, \lambda_{\text{sds}}^0, k_{\text{cool}})$
- 20: $\mathbf{g}_{\text{sds}} \leftarrow \mathbf{g}_{\text{ema}}$
- 21: **else**
- 22: $\mathbf{g}_{\text{sds}} \leftarrow \mathbf{0}; \quad \lambda_{\text{sds}} \leftarrow 0$
- 23: **end if**
- 24: $\mathbf{g} \leftarrow \lambda_{\text{phys}} \cdot \mathbf{g}_{\text{phys}} + \lambda_{\text{sds}} \cdot \mathbf{g}_{\text{sds}}$ {Weighted sum of raw gradients}
- 25: $\rho \leftarrow \text{Adam}(\rho, \mathbf{g}, lr(k))$
- 26: $\rho \leftarrow \text{BisectionProject}(\rho, V^*)$
- 27: $\rho \leftarrow \text{clip}(\rho, 0, 1)$
- 28: **if** $k \bmod N_{\beta} = 0$ **then**
- 29: $\beta \leftarrow \min(2\beta, \beta_{\max})$
- 30: **end if**
- 31: **end for**
- 32: **return** $\hat{\rho}$

4. Experimental setup

The three plate-bending domains (D_{Rect} , D_{Circle} , D_{Hole}) use Mindlin–Reissner plate theory with the MITC4 element (3 DOF/node) at $V^* = 0.20$. The industrial D_{Link} domain uses plane-stress 4-node bilinear elements (2 DOF/node) at $V^* = 0.30$, with fixed supports on the left arc and combined tension–shear loading ($f_x = f_y = 0.01$) on the right. Domain geometries, load/support configurations, and representative baselines appear in Sections 5.1 and 5.3.

Ten text prompts span three semantic categories (Table 4): biomimetic (*Tree*, *Leaf*, *Bone*), engineering (*Truss*, *Grid*, *Honeycomb*), and geometric (*Voronoi*, *Spiderweb*, *Ornamental*, *Diamond*). Every prompt carries the prefix “bold black lines on white background,” which bridges the density field’s visual style to the diffusion model’s training distribution and mitigates source-distribution mismatch [56]. A negative prompt suppresses blurriness, color, 3D cues, and fine details below the filter radius.

Each run uses $N = 1000$ iterations on an NVIDIA RTX 4090 (24 GB VRAM), taking ~ 44 minutes per SDS run, a ~ 3 -minute ($\sim 7\%$) overhead over the pure-SIMP baseline. Full parameter values appear in Table 3.

Table 3

Summary of experimental parameters.

Parameter	Value	Description
Grid resolution	150×150	FE mesh
Volume fraction V^*	0.20 (0.30 for D_{Link})	Target volume
SIMP penalization p	3.0	Material interpolation
Filter radius $r_{\min}$	5.0	Density filter
Max-thickness radius $r_{\max}$	10.0	Local-volume penalty (App. A.4)
Heaviside β	$1.0 \rightarrow 64.0$	Doubling every 50 iter
Optimizer	<i>Adam</i> ($\beta_1 = 0.5$, $\beta_2 = 0.9$)	Gradient descent
Learning rate	$0.04 \rightarrow 0.01$	Cosine annealing
CFG scale ω	100	Guidance strength
$\lambda_{\text{phys}} / \lambda_{\text{sds}}^0$	$1.0 / 10$ (mech D_{Rect} , D_{Circle} , D_{Link})	Gradient weighting
	$1.0 / 50$ (D_{Hole})	
	$1.0 / 100$ (thermoelastic)	
SDS warmup / cooldown	30 / 500 iter	Scheduling
t -annealing	$0.98 \rightarrow 0.50$	Coarse-to-fine
EMA decay α	0.9	Gradient smoothing
Iterations K	1000	Total steps

Table 4

Classification of text prompts by structural characteristics.

	Biomimetic	Engineering	Geometric
Prompts	<i>Tree, Leaf, Bone</i>	<i>Truss, Grid, Honeycomb</i>	<i>Voronoi, Spiderweb, Ornamental, Diamond</i>
Induced pattern	Hierarchical, Fractal	Regular cells, Tessellated	Irregular, Radial

Statistical significance for SDS-guided improvements uses a one-sided binomial test on the five random seeds: the minimum achievable p -value for 5/5 seeds beating the baseline is $p = 1/2^5 = 0.031$. The analysis is complemented by a one-sample effect size, Cohen's $d = (\bar{C}_{\text{SDS}} - C_{\text{base}}) / s_{\text{SDS}}$ measured relative to the baseline, with a median $|d| = 1.47$ across significant combinations (all corresponding to compliance reductions) indicating large practical effects.

5. Results

5.1. SDS-guided optimization discovers superior topologies

A total of 150 experiments (10 prompts $\times$ 3 domains $\times$ 5 random seeds) ran across three plate-bending domains with volume fraction $V^* = 0.20$. Statistical methodology and effect sizes are detailed in Section 4. In brief, the binomial-null $p = 0.031$ threshold for 5/5 wins defines significance, and median Cohen's $|d| = 1.47$ confirms large practical effect. Table 5 consolidates the results, and Fig. 3 shows representative designs.

D_{Rect} (**150 $\times$ 150, 4-corner pinned, center point load**) offers the framework's smallest headroom. SIMP already converges to an efficient solution, so only 5 of 10 prompts achieve significance, led by *Spiderweb* (-11.9%), *Voronoi*, *Bone*, *Leaf*, and *Tree*. *Grid* and *Diamond* ($\approx +6\%$) actively degrade performance: their rigid lattice conflicts with the diagonal load paths and wastes material in low-stress regions.

D_{Circle} (**circular domain, 4-directional pins, center load**) exposes the value of the generative prior. The combination of non-rectangular geometry and multi-directional supports produces a symmetric

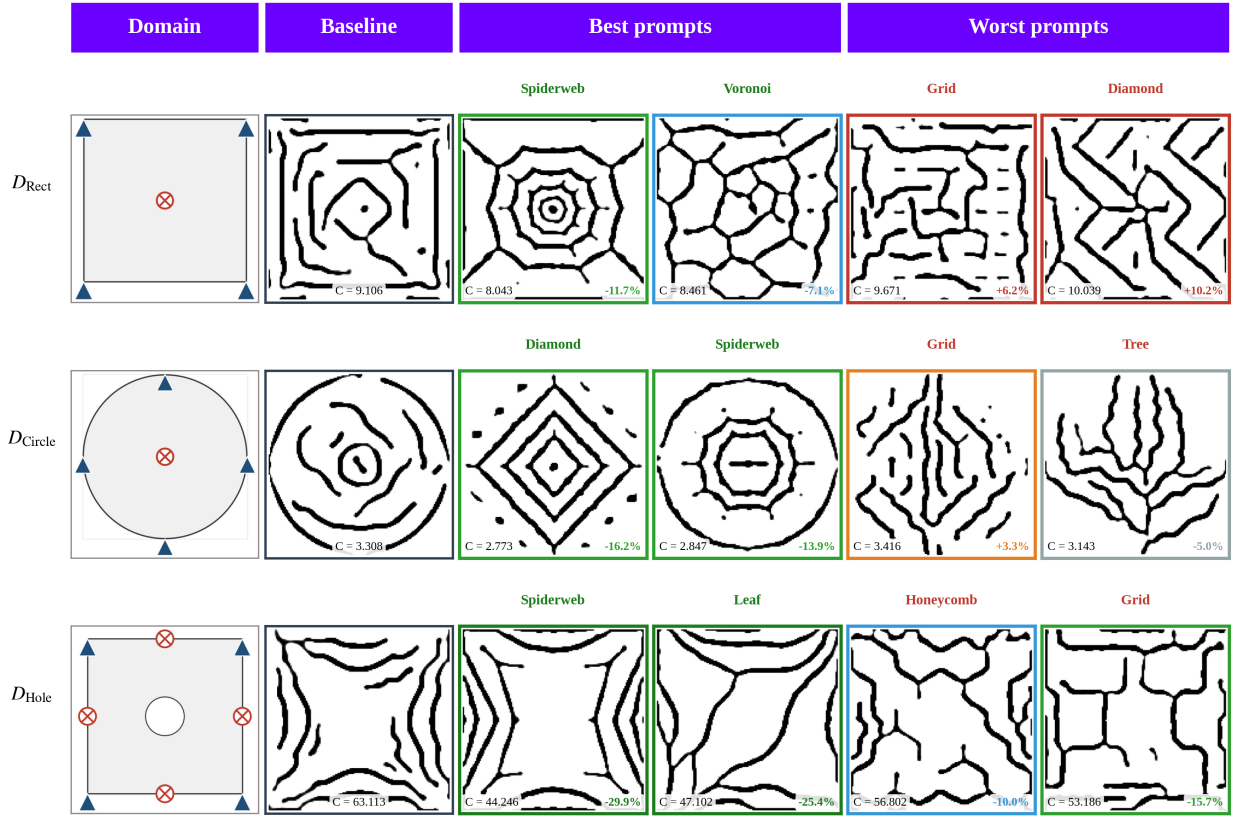


Figure 3: Representative optimized density fields across the three plate-bending domains (D_{Rect} , D_{Circle} , D_{Hole}). Each row: boundary-condition schematic (blue triangles = pin supports; red $\otimes$ = transverse point load) | baseline SIMP | two best-performing and two worst-performing prompts (ranked by 5-seed mean compliance). Each SDS cell displays the seed whose final compliance is closest to the 5-seed mean. Cell-border color encodes $\Delta C/C_{\text{base}}$ (green tones = improvement, red tones = degradation).

landscape with many near-optimal topologies, and SIMP converges to just one. SDS guides the density field to deeper optima that SIMP alone cannot reach, so 9 of 10 prompts achieve significance. *Diamond* leads at -15.2% (best single run -23.8% , $C = 2.522$ vs. baseline 3.308), with *Spiderweb* and *Voronoi* close behind.

D_{Hole} (central hole, 4-side center loads) delivers the framework’s largest gains. The load-path detours around the hole and 4-sided symmetric loading trap the baseline SIMP at $C = 63.113$, roughly seven times the D_{Rect} baseline, so the raw weighted sum calls for a stronger SDS weight ($\lambda_{\text{sds}}^0 = 50$; Section 3.3). Under this setting, all ten prompts achieve 5/5 wins, with *Spiderweb* leading at -28.4% , followed by distributed-branching priors *Leaf* (-25.1%), *Diamond* (-24.6%), and *Bone* (-24.5%). The best single run reaches -31.5% (*Spiderweb*, $C = 43.20$). Tiling priors whose geometry conflicts with the radial load flow lag behind (*Honeycomb* -9.7% , *Grid* -16.7%), foreshadowing the load-path alignment hypothesis developed in Section 5.4. Full mean $\pm$ std per-prompt values appear in Table 5.

5.2. Comparison with alternative baselines

Three alternative SIMP-only baselines test whether simpler methods can replicate the benefits of SDS (Fig. 4). Multi-start SIMP launches optimization from 50 random initial density fields (Gaussian noise, $\sigma = 0.15$) and selects the best result. Although multi-start surpasses the single-seed baseline in all domains, the best SDS result consistently outperforms the multi-start best. Specifically, SDS achieves a best-run compliance of 7.886 versus 8.102 for multi-start in D_{Rect} , 2.522 versus 3.166 in D_{Circle} , and 43.20 versus 47.434 in D_{Hole} (the latter at $\lambda_{\text{sds}}^0 = 50$; Section 5.1; numerical summary in Appendix D.1).

Table 5

Per-prompt compliance C across the four design domains (mean $\pm$ std over $n = 5$ seeds). Bold entries denote statistically significant improvements over the baseline (5/5 seeds win, $p = 0.031$ binomial). The first two rows give the SIMP baseline and the SDS weight λ_{SDS}^0 per domain; the last two rows report the single best run and the number of prompts (out of 10) achieving significance. $V^* = 0.20$ for the three plate-bending domains and $V^* = 0.30$ for the plane-stress D_{Link} . 95% confidence intervals follow as mean $\pm 2.78 \text{ std}/\sqrt{5}$; the per-seed distributions for all combinations are shown in Figs. D.2–D.4.

	D_{Rect}	D_{Circle}	D_{Hole}	D_{Link}
Baseline C_{base}	9.106	3.308	63.113	5.258
λ_{SDS}^0	10	10	50	10
<i>Bone</i>	8.48 $\pm$ 0.07	3.001 $\pm$ 0.14	47.68 $\pm$ 2.42	4.73 $\pm$ 0.18
<i>Diamond</i>	9.65 $\pm$ 0.68	2.804 $\pm$ 0.22	47.59 $\pm$ 2.31	4.80 $\pm$ 0.20
<i>Grid</i>	9.64 $\pm$ 0.06	3.359 $\pm$ 0.14	52.58 $\pm$ 2.76	4.78 $\pm$ 0.21
<i>Honeycomb</i>	9.06 $\pm$ 0.15	3.131 $\pm$ 0.12	56.97 $\pm$ 1.01	4.79 $\pm$ 0.25
<i>Leaf</i>	8.50 $\pm$ 0.09	3.020 $\pm$ 0.07	47.30 $\pm$ 1.14	4.73 $\pm$ 0.19
<i>Ornamental</i>	9.17 $\pm$ 0.35	3.052 $\pm$ 0.06	48.27 $\pm$ 2.82	5.37 $\pm$ 0.75
<i>Spiderweb</i>	8.02 $\pm$ 0.10	2.846 $\pm$ 0.06	45.22 $\pm$ 2.04	5.12 $\pm$ 0.32
<i>Tree</i>	8.51 $\pm$ 0.19	3.143 $\pm$ 0.04	47.78 $\pm$ 0.51	4.81 $\pm$ 0.10
<i>Truss</i>	8.97 $\pm$ 0.18	3.083 $\pm$ 0.05	50.22 $\pm$ 3.55	5.06 $\pm$ 0.18
<i>Voronoi</i>	8.43 $\pm$ 0.12	2.906 $\pm$ 0.05	48.39 $\pm$ 2.67	4.70 $\pm$ 0.10
Best single run C	7.886	2.522	43.20	4.459
Sig. prompts (5/5 wins)	5/10	9/10	10/10	7/10

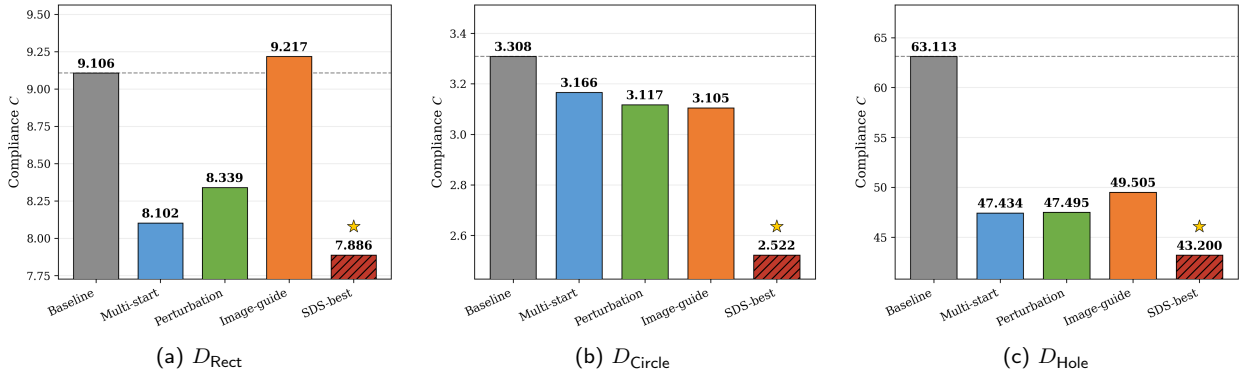


Figure 4: Compliance comparison of SDS-guided optimization against three SIMP-only alternatives (Multi-start, Perturbation, Image-guided) across the three plate-bending domains; the multi-start and perturbation bars are the best of 50 runs per domain, with the matched-number-of-runs distributional analysis given in Appendix D.1.

Random perturbation SIMP periodically injects Gaussian noise into the density field. Such unstructured noise, however, cannot replace the text-conditioned structural guidance of SDS. Consequently, performance remains comparable to or worse than the baseline on average. Similarly, image-guided topology optimization tests whether the benefit originates from the final image or the dynamic optimization process by using the mean squared error loss with a fixed SDS-generated target. While this method partially recovers some benefits in D_{Circle} (3.105) and D_{Hole} (49.505), it remains substantially inferior to the adaptive SDS approach, which achieves 2.522 and 43.20, respectively. In D_{Rect} , image-guided optimization even degrades performance relative to the baseline (9.217 vs. 9.106).

The advantage of SDS therefore lies in its iterative, state-dependent gradient: even when given the exact SDS-generated design as a fixed target, image-guided optimization cannot recover the improvement, because it lacks the adaptive topology reshaping that SDS supplies during intermediate iterations. Multi-start SIMP,

for its part, samples multiple initial conditions but remains limited to the basins of attraction reachable from unstructured Gaussian-noise initializations. In contrast, SDS accesses qualitatively different topological regions by imposing structured, physics-compatible exploration guided by the generative prior. This gap is largest in D_{Hole} , where the complex load paths most severely trap conventional methods. Beyond the conventional OC optimizer, which performs substantially worse (Appendix B.4), a tuned Method of Moving Asymptotes (MMA) optimizer [57], a standard strong baseline in gradient-based topology optimization, is also benchmarked. Run for the same 1000 iterations under identical filtering and β -continuation, MMA improves substantially on the $Adam$ baseline in every domain (Table 6; converged designs in Fig. 5) and thus serves as a strong single-start gradient baseline. Even against this baseline, the best SDS run attains lower compliance in the two multi-directional domains, D_{Circle} (2.522 vs. 2.598) and D_{Hole} (43.20 vs. 48.85). Because MMA is a single deterministic run, this best-run comparison is the matched one; on the five-seed mean the stochastic SDS prompts straddle MMA , as expected for a method with run-to-run variation. It trails only in the single-load radial domain D_{Rect} (7.886 vs. 7.721), where the radial load path makes the compliance landscape simple enough that gradient descent already reaches a near-global optimum. The comparison is conservative in favour of MMA . It is $\approx 2.1\times$ more expensive per iteration than $Adam$ (4.90 vs. 2.34 s/iter), so an equal-iteration comparison already grants it roughly twice the compute budget of the gradient baselines. MMA is also a deterministic single-start method, whereas the matched-number-of-runs multi-start study (Appendix D.1) shows that even 50 random restarts do not reach the SDS result. That SDS still prevails wherever the load paths are non-trivial indicates that its contribution is not to win a compliance race against a tuned optimizer. Steered by a text prompt and without task-specific training, it reaches topological basins that gradient descent does not access, precisely in the domains where the design problem is hard.

Table 6

Baseline compliance: tuned MMA vs. $Adam$ and the best SDS prompt (no SDS for the optimizer baselines; $V^* = 0.20$, 1000 iterations). MMA improves on $Adam$ everywhere but is surpassed by SDS in the two multi-directional domains.

Domain	$Adam$ C	MMA C	SDS best C	Lowest
D_{Rect}	9.106	7.721	7.886	MMA
D_{Circle}	3.308	2.598	2.522	SDS
D_{Hole}	63.113	48.848	43.20	SDS

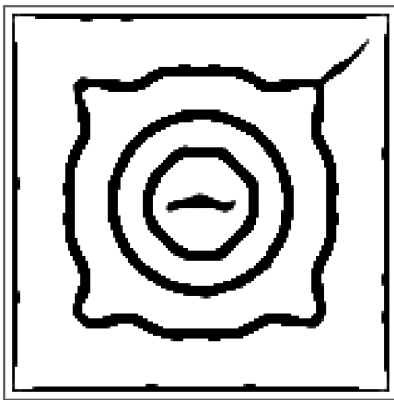
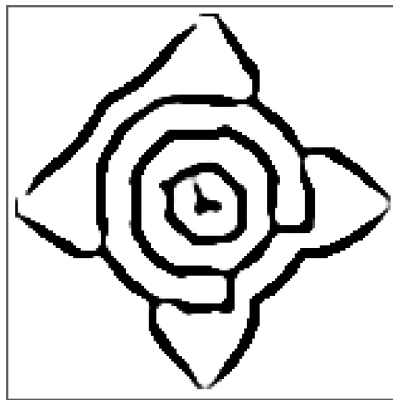
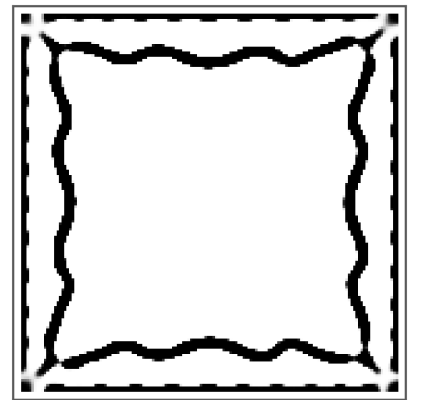
(a) D_{Rect} ($C = 7.721$)(b) D_{Circle} ($C = 2.598$)(c) D_{Hole} ($C = 48.848$)

Figure 5: Converged MMA baseline designs for the three plate-bending domains (the MMA column of Table 6; $V^* = 0.20$, 1000 iterations, Heaviside β -continuation, filter radii $r_{\min} = 5$ and $r_{\max} = 10$). The per-panel compliance C is the best value attained over the run.

The cost contrast with data-driven generative alternatives is also large. Data-driven alternatives such as TopoDiff [21] and DOM [22] require thousands of GPU-hours and large pre-computed solution datasets (Table 1). The present framework eliminates both: retargeting to a new boundary condition, physics objective, or domain geometry reduces to a single optimization run (Section 4) with zero label cost. The saving recurs for each new condition explored, since retargeting needs no retraining.

5.3. Generalization to industrial geometry

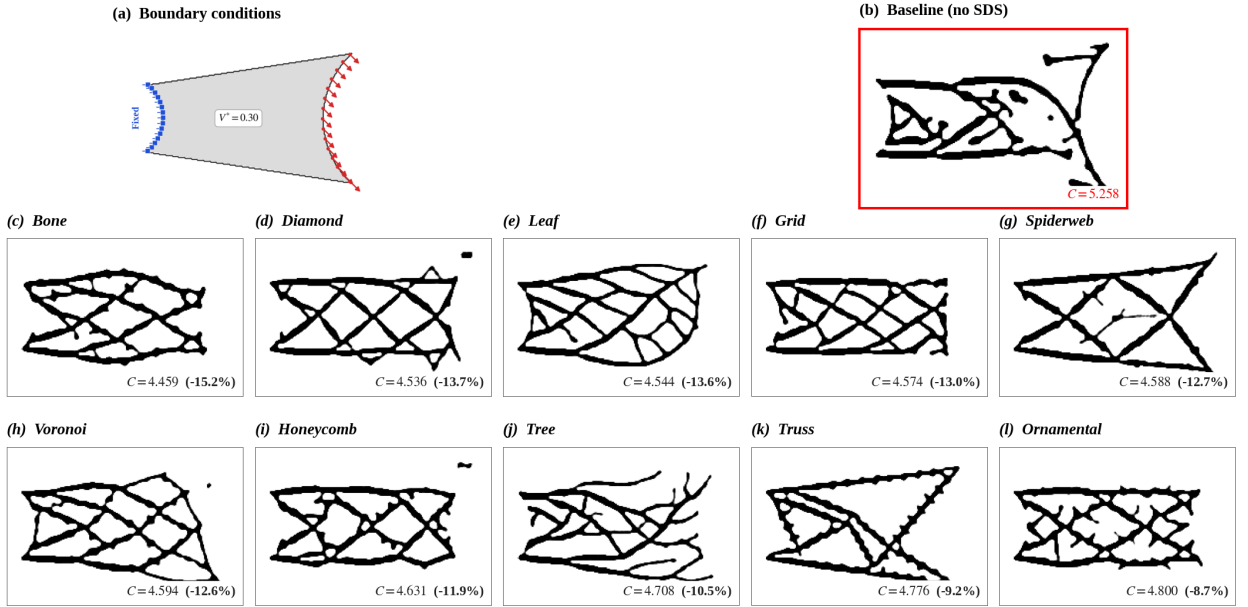


Figure 6: Optimized density fields for the D_{Link} domain (automotive suspension link, plane stress, $V^* = 0.30$). (a) Boundary conditions: the left arc is fixed and the right arc carries combined tension–shear loading; (b) baseline SIMP (no SDS); (c–l) designs under ten text prompts, ordered by ascending compliance. The irregular domain geometry and distributed boundary conditions produce complex, non-radial load paths that are qualitatively distinct from the synthetic domains.

To test generalization beyond synthetic benchmarks, 50 additional experiments (10 prompts $\times$ 5 seeds) ran on D_{Link} , which uses the 2D cross-section of an automotive suspension link as the design domain (Fig. 6). Unlike the three plate-bending domains analyzed under Mindlin–Reissner plate theory, D_{Link} employs plane stress analysis (4-node bilinear element, 2 DOF per node) with $V^* = 0.30$, and features an asymmetric geometry with distributed supports on the left arc and combined tension–shear loading ($f_x = f_y = 0.01$) on the right arc. Seven out of ten prompts achieved statistically significant improvements (5/5 wins, $p = 0.031$). See D_{Link} column of Table 5). The top performer *Voronoi* reduced mean compliance by -10.6% , closely followed by *Leaf* (-10.0%) and *Bone* (-10.0%). These leading prompts share a common characteristic: they induce distributed, irregularly branching networks that can adapt freely to asymmetric left-to-right load transmission paths. Conversely, the only prompt that degraded performance (*Ornamental*, $+2.1\%$) exhibited the highest seed-level variability (std = 0.75), consistent with the observation that misaligned prompts produce unreliable results. Full mean $\pm$ std values per prompt appear in Table 5.

5.4. Prompt–domain interaction governs performance

Prompt rankings across the four domains reveal pronounced reversals that define a *load-path alignment hypothesis*: compliance improves when the SDS-induced pattern geometrically aligns with the domain’s actual load paths (Fig. 7).

Spiderweb, which induces a radial pattern, leads in D_{Rect} (-11.9%) and D_{Hole} (-28.4%) and ranks second in D_{Circle} (-14.0%), where center-to-support load paths naturally align with radial geometry. The same prompt drops to 9th in the industrial D_{Link} domain (-2.6%), because its single-focal-point bias does

Zero-shot rib design

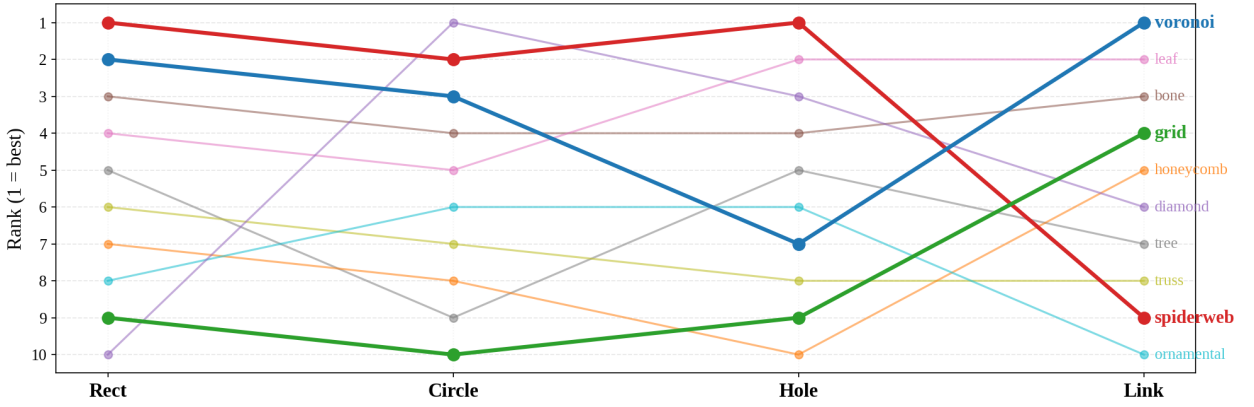


Figure 7: Parallel-coordinates plot of prompt rankings across the four design domains (Rank 1 = lowest mean compliance). Three prompts with the most informative trajectories (*Spiderweb*, *Voronoi*, *Grid*) are highlighted; the remaining seven appear in muted colors.

not match the distributed arc loading. *Voronoi* shows the opposite pattern: its irregular cell partitioning imposes minimal geometric constraint, letting the physics solver adapt to distributed loads, which lifts it from 2nd in D_{Rect} and 3rd in D_{Circle} to 1st in D_{Link} (-10.6%). Biomimetic branching priors (*Leaf*, *Bone*) remain broadly effective, placing 3rd–4th in D_{Rect} and among the top four in D_{Hole} and D_{Link} .

Grid underperforms in every radially loaded domain: $+5.9\%$ degradation in D_{Rect} , 10th in D_{Circle} ($+1.5\%$), and 9th in D_{Hole} (-16.7%). Its orthogonal geometry conflicts with radial load flow. The exception is D_{Link} , where *Grid* rises to 4th (-9.1%) because the orthogonal lattice aligns productively with the tension–shear decomposition of the distributed arc loading. The same alignment principle holds even in the most deeply trapped domain D_{Hole} , where radial/branching priors cluster near -25% while conflicting tiling priors (*Honeycomb* -9.7%) trail the group by more than 15 percentage points.

These rank reversals are a direct consequence of making prompt–domain compatibility explicit through physics feedback at every iteration, rather than learning it implicitly from labeled data as a trained generative model would. The design implications are developed in Section 7.

5.5. Generalization to thermoelastic loading

A further test applies the framework to thermoelastic compliance minimization, where a uniform through-thickness temperature gradient ΔT induces distributed thermal bending moments rather than concentrated external forces [50]. The objective becomes $C = \mathbf{F}_{\text{th}}^T \mathbf{u}$, with the thermal load prescribed (design-independent) so the sensitivity retains its standard self-adjoint form. The MITC4 plate FEA, SIMP interpolation, Heaviside projection, and SDS scheduling carry over unchanged. Only the load vector changes.

The SDS weight scales to $\lambda_{\text{sds}}^0 = 100$ to match the larger and more diffuse thermal physics gradient (Section 3.3; empirical sweep in Appendix D.6). Three prompts (*Spiderweb*, *Voronoi*, *Tree*) ran with five seeds each across D_{Rect} , D_{Circle} , and D_{Hole} . Fig. 8 shows representative designs and Table 7 the full results. In D_{Rect} (baseline $C = 613.6$), all three prompts reached 5/5 wins with mean improvements up to -23.0% . In D_{Circle} (baseline $C = 818.2$), all three reached 5/5 wins (-9.6% to -13.9%). In D_{Hole} (baseline $C = 556.9$), *Voronoi* achieved 5/5 significance (-8.5%) while *Spiderweb* and *Tree* reached 4/5 with -7.2% and -2.6% mean improvement, and no prompt degraded the baseline. Across all thermoelastic settings, 7 of 9 combinations reached 5/5 wins and the remaining two reached 4/5.

The domain difficulty ranking reverses relative to mechanical loading: D_{Rect} , the easiest mechanical domain, becomes the most responsive under uniform thermal loading because its isotropic thermal moment admits many competing near-optimal topologies that the generative prior helps discriminate among. The central hole in D_{Hole} , conversely, geometrically constrains feasible topologies and drives the baseline close to optimality ($C = 556.9$ vs. 613.6 for D_{Rect}), leaving little headroom. This reversal is consistent with the load-path alignment hypothesis: radial priors still succeed at D_{Circle} and branching priors at D_{Rect} , but

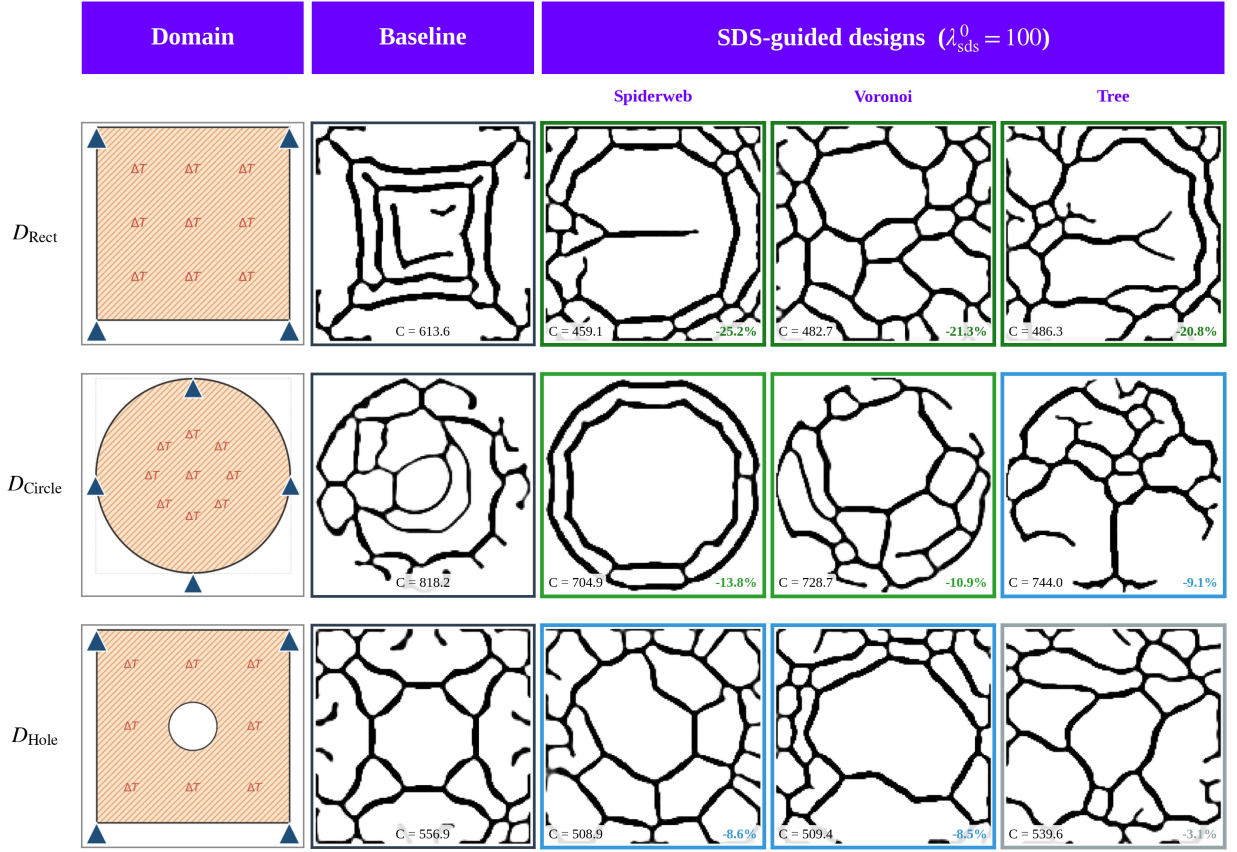


Figure 8: Thermoelastic topology optimization under a uniform through-thickness temperature gradient ΔT . Each row corresponds to one domain (D_{Rect} , D_{Circle} , D_{Hole}); columns show the boundary-condition schematic, the baseline SIMP solution, and SDS-guided designs at $\lambda_{\text{sds}}^0 = 100$ for *Spiderweb*, *Voronoi*, and *Tree*. Cell-border color encodes the improvement using the same scale as Fig. 3; per-cell compliance C and $\Delta\%$ are overlaid.

Table 7

Thermoelastic compliance results with $\lambda_{\text{sds}}^0 = 100$ (mean $\pm$ std over 5 seeds, $V^* = 0.20$). Values are thermoelastic compliance C ; lower is better. Bold with “5/5” indicates statistically significant improvement (5/5 wins vs. baseline); “4/5” denotes four out of five seeds improving over the baseline.

	D_{Rect}	D_{Circle}	D_{Hole}
Baseline (<i>Adam</i>)	613.6	818.2	556.9
<i>Spiderweb</i>	472.5 $\pm$ 37 (−23.0%, 5/5)	704.6 $\pm$ 4 (−13.9%, 5/5)	516.5 $\pm$ 36 (−7.2%, 4/5)
<i>Voronoi</i>	480.5 $\pm$ 6 (−21.7%, 5/5)	727.7 $\pm$ 19 (−11.1%, 5/5)	509.6 $\pm$ 10 (−8.5%, 5/5)
<i>Tree</i>	486.9 $\pm$ 2 (−20.7%, 5/5)	739.9 $\pm$ 15 (−9.6%, 5/5)	542.4 $\pm$ 12 (−2.6%, 4/5)
Best $\Delta\%$	−23.0%	−13.9%	−8.5%

prompts whose spatial bias is less directly mirrored by the loading benefit from stronger generative weight to compete against the larger physics gradient.

6. Mechanisms of improvement

Across the six experiments (mechanical point loads, distributed arc loading, and thermoelastic bending), the same pattern of improvement recurs. This section examines the structural mechanism behind SDS-guided gains, the conditions under which it operates, and the roles of the scheduling and gray-density components.

6.1. Structural signature: dead-end suppression

In most domains tested, SDS-guided designs tend to eliminate non-load-bearing dead-end branches from the rib skeleton (free ends that extend from the skeleton but fail to reach a support). This runs counter to a natural expectation. Because generative image models are tuned for visual diversity, one might anticipate that a diffusion prior would enrich the skeleton with additional branches. Coupled to the FEA sensitivity, however, SDS does the opposite (Fig. 9). In the baseline design (left), many skeletal branches terminate without reaching a support, and the red endpoints in the magnified inset mark ribs that carry essentially no load, whereas the SDS-guided design (right) reroutes or removes these stubs, leaving a sparser network of fully connected load paths. The generative prior therefore simplifies the structure rather than diversifying it, and this pruning of dead ends is what improves physical performance.

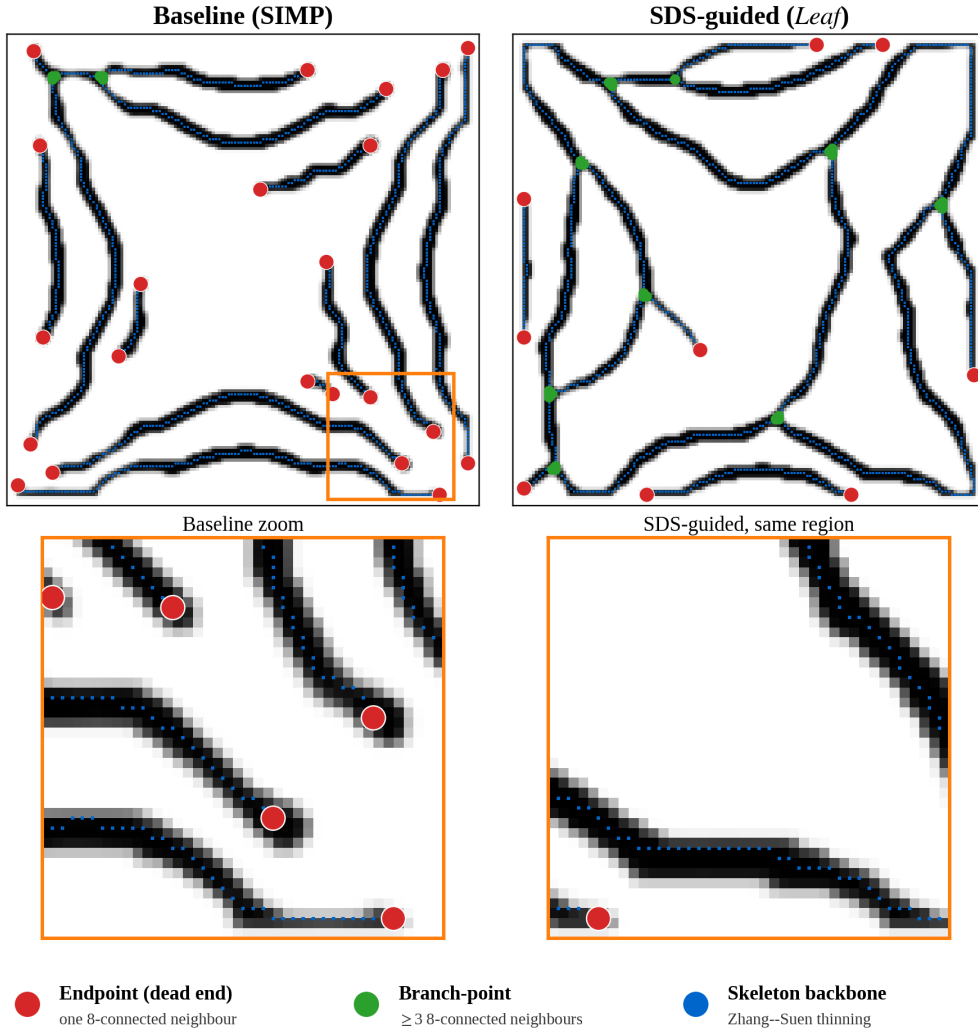


Figure 9: Qualitative illustration of dead-end suppression on D_{Hole} . Skeletons (Zhang–Suen thinning, blue) of the baseline SIMP design (left) and an SDS-guided design (*Leaf*, right), with endpoints/dead ends (red, one 8-connected neighbour) and branch points (green, three or more). The magnified insets show that the baseline retains numerous free-ending ribs that never reach a support, whereas the SDS-guided network reroutes or removes them, simplifying rather than enriching the topology.

Stable Diffusion, trained on billions of image-text pairs from LAION-5B [58], produces continuous linear networks in response to structural prompts prefixed with “bold black lines on white background,” and this

bias propagates through the VAE encoder back to the density field. A natural-image prior helps here for a concrete reason: the thin, branching line networks the model produces for such prompts are the morphologies (leaf venation, trabecular bone, spider webs) that efficient rib layouts approximate (Section 1). The visual bias of the prior and the structural objective are thus morphologically aligned, and the “bold black lines on white background” prefix explicitly steers the model toward this line-network sub-manifold of its training distribution. Morphological analysis across all six experimental domains (Fig. 11 and Table 9) shows that the prompt-level Pearson correlation between mean compliance and skeleton endpoint count n_{end} is positive in every domain ($r = +0.56$ to $+0.99$, median $r = +0.97$). In several domains, top-performing prompts reduce endpoint counts by 30–50% while preserving or extending total skeleton length: *Spiderweb* drops endpoints in D_{Rect} from 38 to 19 and elongates the network from 1090 to 1299 skeleton pixels. *Leaf* in D_{Hole} reduces endpoints from 22 to 13. *Spiderweb* in thermoelastic D_{Circle} collapses endpoints from 11 to below 1. By contrast, neither branch-point count n_{branch} nor raw connectivity correlates consistently with compliance across domains. The dot-prompt ablation provides controlled evidence consistent with this mechanism (Fig. 10; setup and full table in Appendix D.5). On D_{Rect} , changing a single prompt word (“lines”→“dots”) so that the prior places isolated, non-load-bearing points inflates the endpoint count beyond any other prompt and yields the largest degradation below the baseline (+7.3%, the only condition with no seed improving on the baseline). The semantically irrelevant *sunset* prompt, with the next-highest endpoint count, is the only other condition to worsen the baseline (+3.6%), so the two most endpoint-rich prompts are exactly the two that degrade compliance. Across the thirty individual runs (five seeds for each of the six prompts) the endpoint count and the compliance change track each other at $r = 0.87$. Semantically empty or prefix-only prompts yield far smaller gains (−3.6% and −4.5%) than the structural prompt (−11.9%). The improvement tracks the endpoint count, which structural prompts reliably minimize, rather than the mere presence of generative coupling. Even the semantically irrelevant *fruit* prompt helps (−5.5%) when it happens to yield a low-endpoint morphology. A complementary within-design test isolates the endpoint effect without altering the surrounding morphology: on fixed optimized designs, surgically removing all dead-end twigs (skeleton-endpoint pruning followed by re-analysis) raises compliance by only $\sim 6\%$ on average across nine designs, whereas removing a single load-bearing rib raises it by +61.5% on average (up to +182%), and grafting an artificial dead-end changes compliance by under 1% (Table 8). Dead ends therefore carry comparatively little load, providing direct intervention-based support for the mechanism, more directly than the dot ablation, which co-varies the global morphology. The interventions and per-design values are detailed in Appendix D.9.

Table 8

Within-design causal surgery: mean compliance change (ΔC) over three designs per domain for each intervention. Dead-end removal costs far less than removing a single load-bearing rib, and grafting a dead-end is inert.

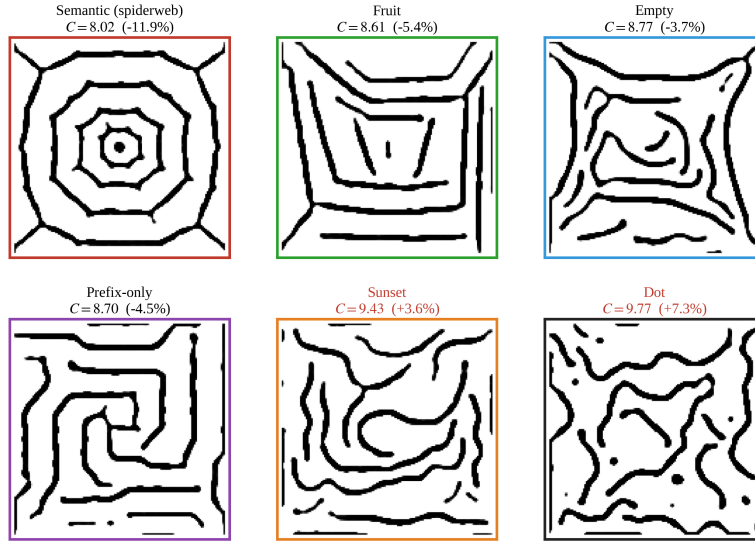
Domain	Prune dead-ends	Prune load-bearing rib	Graft dead-end
D_{Rect}	+6.9%	+18.1%	−0.8%
D_{Circle}	+7.2%	+24.3%	−0.4%
D_{Hole}	+3.5%	+142.0%	−0.1%
All (9 designs)	+5.9%	+61.5%	≈ 0

In most domains, therefore, SDS improves designs by eliminating non-load-bearing dead ends rather than by adding branches. The physical rationale is direct: a dead-end branch, disconnected from any support pathway, carries near-zero stress and contributes negligibly to stiffness, yet consumes volume fraction that the bisection constraint must redistribute from load-carrying members elsewhere. Each eliminated dead end therefore frees material budget for primary load paths without violating the volume target, explaining why endpoint count (not branch count or connectivity alone) tracks compliance across diverse physics settings.

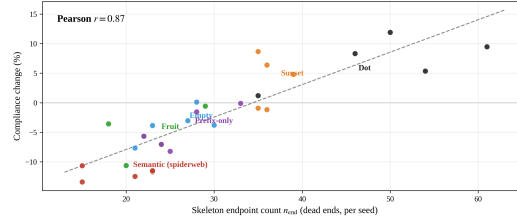
6.2. Physical explanation: escape from local optima

The convergence histories (Fig. 12) make the optimization-level consequence of dead-end suppression directly visible. The SDS gradient deforms the effective energy landscape by adding a text-conditioned potential to the physics-based objective, enabling transitions to basins inaccessible under the physical

Zero-shot rib design

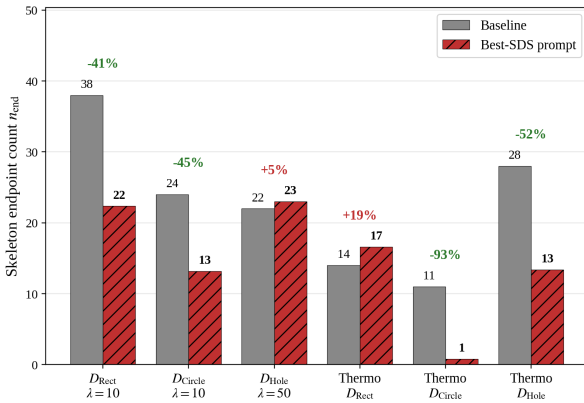


(a) Optimized rib layouts under the six prompt conditions

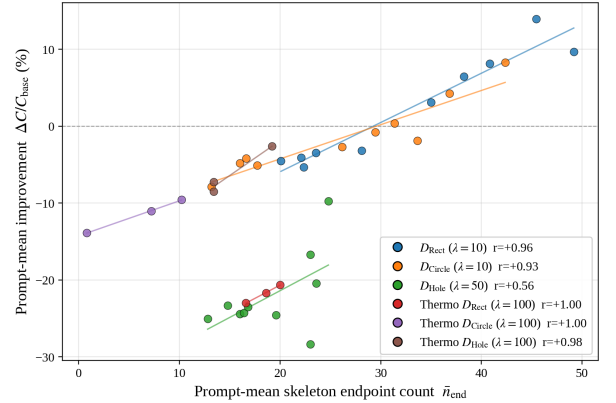


(b) Endpoint count vs. compliance change

Figure 10: Prompt comparison on D_{Rect} . (a) Optimized rib layouts under six prompts (one representative seed each; the labeled C is the five-seed mean, with *Spiderweb* the main-experiment reference), ordered from the best structural prompt (*Spiderweb*) to the *dot* prompt: as the prompt shifts from connected line networks toward isolated points, the layout accumulates fragmented, non-load-bearing stubs. (b) Per-seed compliance change versus per-seed skeleton endpoint count n_{end} (five seeds for each of the six prompts, 30 runs); the two correlate at Pearson $r = 0.87$, and the two highest- n_{end} conditions (*dot* and *sunset*) are the ones that degrade the baseline, the *dot* prompt most.



(a) Baseline vs. best-SDS prompt skeleton endpoint counts across the six experimental domains. Per-bar labels give the percentage change from baseline.



(b) Prompt-level scatter of mean compliance improvement against mean skeleton endpoint count $\bar{n}_{\text{end}}$ across all six domains. Per-domain Pearson r is annotated.

Figure 11: Cross-domain evidence that dead-end suppression is the structural signature of SDS-guided improvement. Each panel shows one representative design per domain, so its endpoint counts differ slightly from the five-seed means in Table 9.

landscape alone, analogous to how DreamFusion [20] discovers meaningful 3D structures from random NeRF initialization. During the warmup phase all curves track each other while SIMP establishes initial load paths. Once SDS activates, the best-prompt curve separates downward from the baseline, locking in the gap by the end of the SDS-active phase. The clearest example is D_{Circle} , where the best prompt reaches a mean $C = 2.804$ (−15.2%) compared to the baseline local optimum at $C = 3.308$: a topology that SIMP from uniform initialization never reaches.

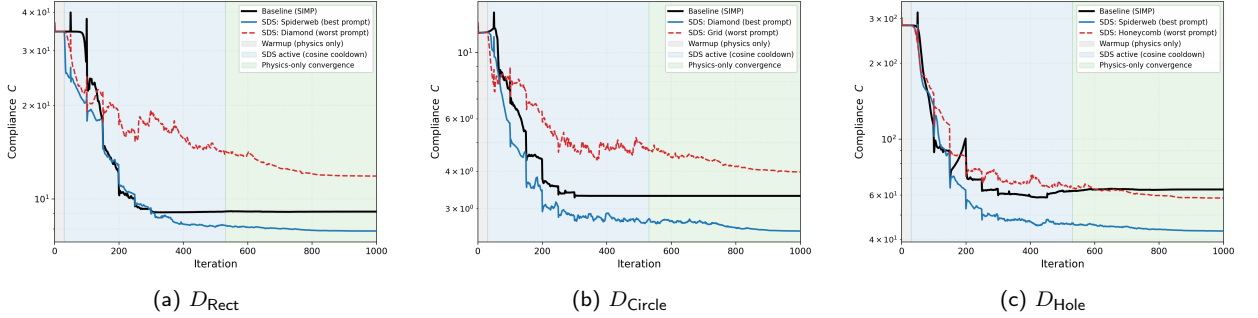


Figure 12: Compliance convergence history on a log-scale y -axis for the three plate-bending domains. Each panel superimposes the baseline SIMP trajectory (black), the best-prompt best-seed run (blue), and a high-compliance worst-prompt run (red dashed). Background shading marks the three scheduling phases: warmup (gray, 0–30), SDS-active cosine cooldown (blue, 30–530), and physics-only convergence (green, 530–1,000).

6.3. Boundary condition: load-path alignment

Dead-end suppression operates most effectively when the prompt’s geometric bias aligns with the domain’s dominant load paths, the condition that governs which prompts succeed in which domains, as detailed in Section 5.4. When prompt and load path mismatch, the generative gradient opposes the physics gradient throughout the SDS-active phase, producing the +6% degradations seen for *Diamond* and *Grid* in D_{Rect} (failure-mode analysis in Appendix D.4). The two mechanisms are therefore complementary: dead-end suppression is the common operation. Load-path alignment determines its effectiveness.

6.4. Resolving the gray-density tendency

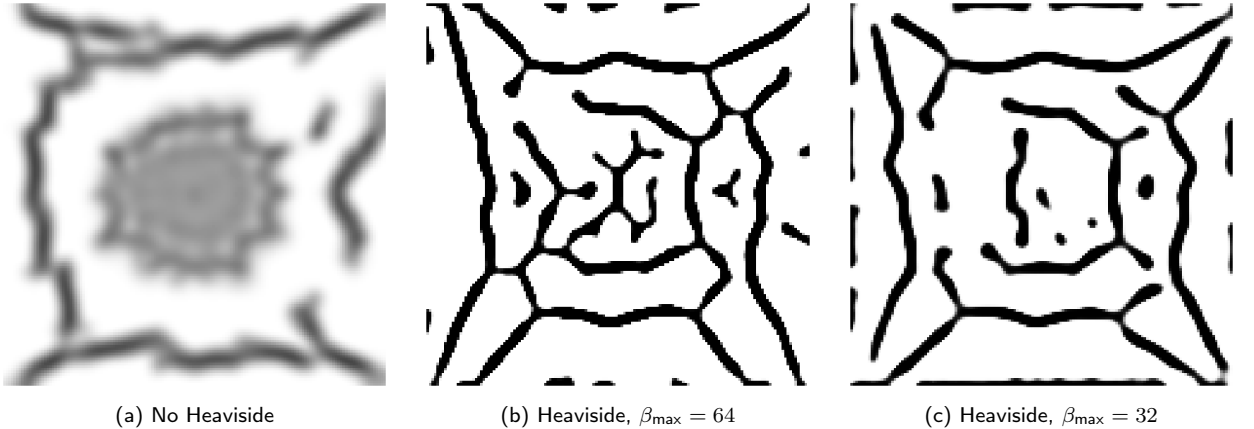


Figure 13: Effect of Heaviside projection and β -continuation on binarization. Without Heaviside (a) the gray-density ratio reaches 42.6%; with $\beta_{\text{max}} = 64$ (b) it drops below 3%.

Table 9

Network morphology metrics of representative designs across all six experimental domains. Values are per-prompt means over $n = 5$ seeds of (i) final compliance C , (ii) skeleton endpoint count n_{end} (non-load-bearing terminal branches, i.e. dead ends), (iii) branch-point count n_{branch} , and (iv) total skeleton length in pixel units. n_{end} correlates positively with C within every domain tested (Pearson r reported in the last column), whereas n_{branch} shows no consistent sign across domains. Rows within each domain are ordered by final compliance (best first); the last row per domain is the worst-performing prompt.

Domain	Prompt	C	n_{end}	n_{branch}	ℓ_{skel}	$r(C, n_{\text{end}})$
$D_{\text{Rect}} (\lambda_{\text{sds}}^0 = 10)$	Baseline	9.106	38.0	8.0	1090	+0.96
	<i>Spiderweb</i>	8.022	19.2	40.6	1299	
	<i>Voronoi</i>	8.433	14.8	98.2	1179	
	<i>Grid</i>	9.643	52.0	17.2	1040	
$D_{\text{Circle}} (\lambda_{\text{sds}}^0 = 10)$	Baseline	3.308	24.0	0.0	765	+0.93
	<i>Diamond</i>	2.804	24.0	31.2	675	
	<i>Spiderweb</i>	2.846	14.4	22.4	833	
	<i>Grid</i>	3.359	32.2	26.2	727	
$D_{\text{Hole}} (\lambda_{\text{sds}}^0 = 50)$	Baseline	63.113	22.0	6.0	978	+0.56
	<i>Spiderweb</i>	45.22	23.0	18.8	1100	
	<i>Leaf</i>	47.30	12.8	23.8	1048	
	<i>Honeycomb</i>	56.97	24.8	42.8	1006	
Thermo $D_{\text{Rect}} (\lambda_{\text{sds}}^0 = 100)$	Baseline	613.6	14.0	48.0	1056	+0.99
	<i>Spiderweb</i>	472.5	16.6	121.8	1242	
	<i>Voronoi</i>	480.5	18.6	130.6	1204	
	<i>Tree</i>	486.9	20.0	111.0	1196	
Thermo $D_{\text{Circle}} (\lambda_{\text{sds}}^0 = 100)$	Baseline	818.2	11.0	88.0	1019	+0.99
	<i>Spiderweb</i>	704.6	0.8	24.0	742	
	<i>Voronoi</i>	727.7	7.2	79.6	871	
	<i>Tree</i>	739.9	10.2	72.2	828	
Thermo $D_{\text{Hole}} (\lambda_{\text{sds}}^0 = 100)$	Baseline	556.9	28.0	90.0	1087	+0.98
	<i>Voronoi</i>	509.6	13.4	87.6	1079	
	<i>Spiderweb</i>	516.5	13.4	85.4	1074	
	<i>Tree</i>	542.4	19.2	92.8	1067	

SDS-SIMP coupling inherits a distinct manufacturability challenge: the score function of Stable Diffusion operates in continuous image pixel space, and the smoothness of its training distribution biases the density field toward intermediate values. Without correction, the gray-density ratio of SDS-guided optimization reaches 42.6%, nearly double the $\sim 20\%$ of pure SIMP (Fig. 13). This did not affect DreamFusion’s NeRF outputs (which are inherently continuous radiance fields), but it renders SDS-guided density fields physically uninterpretable.

Heaviside projection with β -continuation [46, 59] resolves the issue. The projection and continuation schemes are themselves standard [46, 59]. What is specific to this work is the identification that SDS’s continuous pixel-space bias roughly doubles the gray-density ratio relative to pure SIMP, and the coupling of the β -schedule to the SDS schedule so that generative exploration and binarization do not interfere. Doubling β every 50 iterations from $\beta_0 = 1$ to $\beta_{\text{max}} = 64$ permits smooth global exploration early and enforces sharp binarization later, reducing the gray ratio to 2.5% at negligible compliance cost (quantitative sweep over β_{max} in Appendix D.3). The β -schedule synergizes with SDS scheduling: low- β warmup coincides with early SDS activity to allow generative gradients to propagate across the entire field, while high- β (64) phases overlap with SDS cooldown so the physics gradient finalizes the binarized design.

6.5. Scheduling component ablation

To quantitatively verify the contribution of each component in the four-stage SDS scheduling strategy, a one-at-a-time ablation study was conducted under the *Spiderweb* $\times$ D_{Rect} $\times$ (seed = 0) condition, selected because it exhibits the most consistent improvement (-11.9%) across multiple seeds. Table 10 presents the complete results, including the thermoelastic λ_{sds}^0 sweep used to set the thermal SDS weight.

Table 10

SDS scheduling ablation (*Spiderweb* prompt). The upper block perturbs scheduling hyperparameters in the mechanical D_{Rect} setting ($\lambda_{\text{sds}}^0 = 10$ default, seed 0, default-SDS run $C = 7.91$); the lower block sweeps λ_{sds}^0 in the thermoelastic D_{Rect} setting ($n = 5$ seeds, baseline $C = 613.6$).

Setting	Compliance	vs. Default (mech) / baseline (thermo)
Mechanical D_{Rect} ($\lambda_{\text{sds}}^0 = 10$ default, seed 0)		
Default	7.91	—
<i>Weighting</i>		
$\lambda_{\text{sds}} = 1$	9.60	+21.4%
$\lambda_{\text{sds}} = 5$	8.38	+6.0%
$\lambda_{\text{sds}} = 20$	8.45	+6.9%
<i>Scheduling</i>		
No t -annealing	8.94	+13.0%
warmup = 0	8.57	+8.3%
warmup = 100	7.98	+0.9%
<i>Guidance</i>		
$\omega = 50$	8.28	+4.7%
$\omega = 10$	8.57	+8.4%
$\omega = 200$	8.67	+9.7%
<i>Smoothing</i>		
$\alpha = 0.5$	8.18	+3.4%
$\alpha = 0.0$ (off)	8.31	+5.0%
Thermoelastic D_{Rect} (λ_{sds}^0 sweep, mean over 5 seeds)		
$\lambda_{\text{sds}}^0 = 10$	507.6	−17.3%
$\lambda_{\text{sds}}^0 = 50$	486.6	−20.7%
$\lambda_{\text{sds}}^0 = 100$	472.5	− 23.0%
$\lambda_{\text{sds}}^0 = 200$	490.9	−20.0%

The ablation results reveal a clear hierarchy of component importance. Gradient weighting (λ_{sds}) has the largest impact (+6.0–21.4%): a U-shaped relationship is observed, in which $\lambda_{\text{sds}} = 1$ causes +21.4% degradation due to insufficient SDS influence, while $\lambda_{\text{sds}} = 20$ causes +6.9% degradation from SDS overwhelming the physics gradient. The optimal $\lambda_{\text{sds}} = 10$ balances the two gradients, as the SDS gradient norm is experimentally approximately 0.1 times that of the physics gradient. This confirms that SDS is an auxiliary to physics-based optimization, not a replacement [38, 39]. Timestep annealing is the second most influential factor (+13.0%): without annealing, random timesteps are sampled uniformly throughout optimization, leading to inefficient guidance in which local details are induced before global structure is established. CFG scale (ω) shows +4.7–9.7% sensitivity, with both $\omega = 10$ (insufficient text guidance) and $\omega = 200$ (over-saturation causing gradient instability) degrading performance. $\omega = 100$ matches the original DreamFusion setting [20]. EMA smoothing (α) contributes +3.4–5.0%, and warmup +0.9–8.3%: immediate SDS activation (warmup = 0) biases the uniform initial density field in physically meaningless directions, while excessive warmup (100 iterations) delays SDS engagement to a point where β is already elevated.

In summary, every single-component modification degrades performance, indicating that each scheduling component addresses a distinct failure mode rather than providing redundant regularization. Gradient weight scheduling and timestep annealing serve as the primary coupling mechanisms between SDS and SIMP, EMA provides stability, and Heaviside projection is indispensable for binarization. The thermoelastic λ_{sds}^0 sweep (lower block of Table 10) mirrors the mechanical U-shape but is shifted roughly tenfold toward larger values, quantifying the claim that the optimal SDS weight scales with the magnitude of the physics gradient under the raw weighted sum (Appendix D.6).

7. Discussion

The central finding of this work is that engineer intent, expressed as a natural-language prompt, changes the outcome of a physics-constrained optimization by statistically significant margins. The compliance reductions reported in Section 5 are not stochastic by-products of perturbing the optimizer: in most experimental settings they trace to one operation, the suppression of non-load-bearing dead ends in the rib skeleton (Section 6).

The methodological contribution is the active gradient-level coupling of a pretrained visual prior to a finite-element sensitivity at every iteration. Prior generative approaches to topology optimization either train a domain-specific model on thousands of labeled solutions [21–23] or generate designs independently of physical analysis, so their outputs may be visually plausible yet physically suboptimal. The framework introduced here keeps the diffusion model frozen and lets the FEA gradient, not a loss function trained on examples, control what survives the update. This sets SDS-guided topology optimization apart from prior text- and data-driven approaches. Data-driven generative design needs labeled examples, and CLIP-based text guidance [41] acts only through a global image–text similarity without pixel-level gradients. Semantic, preference-guided large-model optimization [43] works at a complementary, higher level. Here, by contrast, the generative prior contributes pixel-level gradients directly to each density update.

This broadens how design intent can be specified. Conventional parametric approaches encode intent through pre-chosen bases (rib orientations, truss templates, lattice cells), which implicitly restrict the design space to combinations the designer foresaw. For instance, a rib parameterization over four orientation angles, three spacings, and two branching styles enumerates only ~ 24 distinct topologies. A prompt such as “leaf venation,” by contrast, addresses the entire manifold of leaf-like patterns embedded in a pretrained diffusion model’s billion-image prior, with no human enumeration required. The FEA gradient continuously filters these prompt-induced candidates against physical optimality, so engineer intent becomes a direct optimization input rather than a pre-commitment that forecloses exploration. Which regions of this expanded space actually pay off is governed by load-path alignment, as the cross-domain rank reversals make clear: *Spiderweb* is excellent in $D_{\text{Rect}}/D_{\text{Circle}}/D_{\text{Hole}}$ yet drops to 9th in D_{Link} , and *Voronoi* rises from 3rd to 1st between D_{Circle} and D_{Link} . The framework surfaces this compatibility through physics feedback at every iteration, so the expansion is neither uniform nor automatic. This exposes the central assumption behind using a natural-image prior for structural design: nothing guarantees that a visual prior aligns with a given load path. The results make the failure mode explicit. When the prompt-induced morphology conflicts with the dominant load paths, the generative gradient opposes the physics gradient and can degrade compliance (+6% for *Diamond/Grid* in D_{Rect} ; see Sections 5.4 and 6.3). The framework stays robust because the FEA sensitivity continuously constrains the prompt-induced update, and the scheduled decay of the SDS weight ensures that the final optimization stage is governed by the physics gradient alone. Cross-domain transfer of the visual prior is therefore beneficial only insofar as physics filters it, which is also why prompt screening remains necessary in practice.

For the broader AI-for-science and knowledge-intensive design communities, the work extends SDS beyond its native vision domain into physics-constrained optimization, where the output must satisfy equilibrium and volume constraints rather than achieve visual plausibility. Two transferable techniques emerge. First, the optimal SDS weight λ_{SDS}^0 scales with the magnitude of the physics gradient under the raw weighted sum, so identical prompt sets generalize across physics regimes once λ_{SDS}^0 is retuned (empirically 10 for single-point loads, ~ 50 for multi-point geometries with elevated baseline compliance, ~ 100 for spatially diffuse thermal loads). Second, β -continuation, developed here to reconcile SDS’s continuous pixel-space bias with the binary nature of density fields, addresses a general problem in AI-guided optimization: reconciling continuous generative priors with discrete or binary solution spaces. The same β -continuation should apply wherever a continuous generative prior must be reconciled with a binary design field, though this is not tested here.

The framework’s benefit scales with the non-convexity of the physics landscape. To probe this boundary, the identical pipeline was applied to 2D heat conduction topology optimization (thermal compliance $C = \mathbf{F}^\top \mathbf{T}$, conductivity interpolation $k_e = k_{\min} + \hat{\rho}_e^p(k_0 - k_{\min})$ with $k_0 = 1.0$, $k_{\min} = 10^{-3}$ [5]) on a representative D_{Rect} configuration: a center point source with $T = 0$ on all four edges, three prompts $\times$ five seeds. SDS guidance either matched or worsened thermal compliance relative to the *Adam* baseline

across all prompts (Table 11). This negative result is consistent with the central finding of this study: heat conduction under SIMP penalization produces a fundamentally smoother objective landscape than structural compliance, and the optimal topology is typically a branching tree whose global structure is largely determined by the source–sink geometry, leaving few alternative topologies for SDS to explore. The baseline SIMP optimizer already converges to near-optimal branching networks, and the SDS gradient introduces perturbations that disrupt rather than improve these solutions. Pretrained generative priors are therefore most valuable for physics problems where gradient-based optimizers are known to trap severely, such as bending-dominated compliance, fluid–structure coupling, photonic crystal design, or metamaterial unit-cell optimization.

Table 11

Heat conduction results: SDS does not improve thermal compliance. Values are thermal compliance C on the D_{Rect} center-source configuration (mean $\pm$ std over 5 seeds); lower is better.

	Baseline (<i>Adam</i>)	<i>Spiderweb</i>	<i>Voronoi</i>	<i>Tree</i>
C	3.547	3.718 ± 0.44	3.969 ± 0.07	3.973 ± 0.15

Two practical limitations bound the present study. First, the framework is 2D: the density-to-image encoding feeds a 2D diffusion prior, so a 3D extension must first resolve the dimensionality mismatch between 3D density fields and 2D priors, with attendant discretization artifacts such as stair-stepping. Second, prompt selection is manual, and the cross-domain rank reversals confirm that no single prompt is universally optimal. The latter motivates a practical screening protocol for deployment: when retargeting to a new domain, one may try three to five prompts from distinct structural categories, one radial (*Spiderweb*), one distributed-branching (*Leaf* or *Bone*), and one tiling (*Grid* or *Honeycomb*), with single-seed pilot runs, promoting only the shortlisted prompts to a full five-seed evaluation. The present study instead evaluates all prompts at five seeds to characterize the complete prompt–domain landscape. Strong prompt–load-path alignment is typically visible within the first ~ 200 iterations of the SDS-active phase. The compliance curve of a well-matched prompt separates downward from the baseline before $k \approx 230$, whereas poorly matched prompts stagnate or drift upward by the same point. This early signal, combined with the U-shaped λ_{SDS}^0 sweep (Table 10; Appendix D.6), reduces the search cost of adapting the framework to an unfamiliar physics regime to a handful of screening runs.

8. Conclusions

This paper presented a zero-shot rib design framework that merges a training-free generative prior with density-based topology optimization. A frozen Stable Diffusion 2.1 is coupled to a Mindlin–Reissner plate (and plane-stress) SIMP solver through score distillation sampling, so that a natural-language prompt acts as an explicit knowledge representation of engineer design intent whose physical consequences are arbitrated by the FEA sensitivity at every iteration. Four-stage scheduling (warmup, cosine cooldown, timestep annealing, EMA smoothing) stabilizes the coupling. Heaviside projection with β -continuation reduces the SDS-induced gray-density ratio from 42.6% to below 3%, and an automated skeleton-based pipeline converts converged density fields into CAD-convertible candidate STEP geometry in seconds.

Across 245 primary SDS runs spanning four geometric domains and two physics regimes, 38 of 49 prompt–domain combinations achieved statistically significant compliance reductions that outperform multi-start, perturbation, and image-guided baselines (Section 5). These 38 combinations remain significant under a Benjamini–Hochberg false-discovery-rate correction across all 49 tests ($q < 0.05$), and carry a median Cohen’s $|d| = 1.47$. These gains trace, in most domains, to a physically interpretable signature (dead-end suppression in the rib skeleton) and a load-path alignment hypothesis explains which prompts succeed in which domains, while a negative result on heat conduction delineates the scope: pretrained generative priors help most where the physics landscape traps gradient-based optimizers severely.

Future work will extend the framework to 3D via multi-view SDS or native 3D diffusion priors [60, 61], automate prior discovery through optimization in the continuous text-embedding space (textual inversion), and explore physics regimes with severe local-optima trapping such as fluid–structure interaction and

metamaterial unit-cell design. More broadly, this work uses a pretrained generative model as a design interface rather than a design generator: coupling a text prompt to a physics-constrained optimizer produces rib layouts that parametric methods and data-driven generative models do not reach.

Declaration of competing interest

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

Funding

This work was supported by grants from the Ministry of Science and ICT (GTL24033-000, N10250154, and No. 2022-0-00986), the Ministry of Trade, Industry and Energy (RS-2025-02317327 and RS-2025-25444634), the Ministry of Oceans and Fisheries (PET0050), and Korea Hydro & Nuclear Power Co., Ltd. (No. 8-Tech-07).

Data availability

No training datasets were generated or used in this study. The proposed framework is training-free and relies on a publicly available pretrained diffusion model (Stable Diffusion 2.1, <https://huggingface.co/stabilityai/stable-diffusion-2-1>). All numerical results reported in this paper are computational outputs of the framework, fully determined by the code, input specifications (text prompts, domain geometries, boundary conditions), and random seeds documented in the main text and appendices. The implementation code developed for this study will be made available by the corresponding author upon reasonable request for non-commercial research use.

CRedit authorship contribution statement

Yongmin Kwon: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft. **Namwoo Kang:** Project administration, Supervision, Investigation, Funding acquisition, Writing – review & editing.

A. Finite element and SIMP modeling details

A.1. Mindlin–Reissner plate FEA

The Mindlin–Reissner plate theory [47, 48] is adopted to analyze the bending behavior of rib-reinforced plates. Since rib-reinforced plates behave locally as thick plates in the rib regions, the Mindlin–Reissner theory, which accounts for transverse shear deformation, is more appropriate than the Kirchhoff plate theory, which neglects shear deformation. The displacement field is defined as follows:

$$u(x, y, z) = z \cdot \theta_y(x, y), \quad v(x, y, z) = -z \cdot \theta_x(x, y), \quad w = w(x, y) \quad (\text{A.1})$$

where w is the transverse displacement, θ_x the rotation about the x -axis, and θ_y the rotation about the y -axis. The bending strains and shear strains are given respectively as:

$$\boldsymbol{\kappa} = \begin{pmatrix} \partial\theta_y/\partial x \\ -\partial\theta_x/\partial y \\ -\partial\theta_x/\partial x + \partial\theta_y/\partial y \end{pmatrix}, \quad \boldsymbol{\gamma} = \begin{pmatrix} \partial w/\partial x + \theta_y \\ \partial w/\partial y - \theta_x \end{pmatrix} \quad (\text{A.2})$$

In the Kirchhoff theory, $\boldsymbol{\gamma} = \mathbf{0}$ is enforced, whereas the Mindlin–Reissner theory permits $\boldsymbol{\gamma} \neq \mathbf{0}$, thereby accurately capturing the shear strain energy in thick plates. By explicitly accounting for transverse shear deformation ($\theta_x \neq \partial w/\partial x$, $\theta_y \neq \partial w/\partial y$), the Mindlin–Reissner theory is well suited for structures with localized stiffness variations such as ribs.

To numerically implement this theory without shear locking, the MITC4 element is employed [49]. This 4-node bilinear quadrilateral element has three degrees of freedom per node (w, θ_x, θ_y), and the shape

functions are $N_i(\xi, \eta) = \frac{1}{4}(1 + \xi_i\xi)(1 + \eta_i\eta)$. The element stiffness matrix is decomposed into bending and shear contributions as given in Eq. (2), where $\mathbf{D}_b$ is the bending constitutive matrix (3×3) and $\mathbf{D}_s$ is the shear constitutive matrix (2×2 , including the Reissner correction factor $\kappa_s = 5/6$). A standard displacement-based 4-node Mindlin element suffers from shear locking in thin plates due to parasitic shear energy. The MITC4 element overcomes this by using an assumed strain technique that evaluates shear strains at four tying points located at the midpoints of the element edges and then interpolates them into the interior, yielding accurate solutions even in the thin-plate limit ($h \rightarrow 0$).

A.2. SIMP interpolation: rib stiffness modeling

In conventional SIMP, the elastic modulus is interpolated as $E_e = \rho_e^p E_0$, effectively reducing the stiffness of zero-density regions to nearly zero. However, in the rib-reinforced plate model, zero-density regions still have physical substance, as in the base plate. Therefore, a modified SIMP scheme is adopted that directly interpolates bending and shear stiffnesses based on the presence or absence of ribs. Regions with attached ribs have an effective thickness $t_{\max}$ (base plate thickness + rib height), while regions without ribs have only the base plate thickness $t_{\min}$. Since the flexural rigidity of a plate is proportional to the cube of the thickness ($D_b \propto t^3$) and the shear stiffness is linearly proportional to the thickness ($D_s \propto t$), the stiffness varies significantly depending on the presence or absence of ribs. To reflect this physical difference, the stiffness matrices for the two limiting states are defined as follows:

$$\mathbf{D}_b^{\max} = \frac{Et_{\max}^3}{12(1-\nu^2)} \mathbf{C}_b, \quad \mathbf{D}_b^{\min} = \frac{Et_{\min}^3}{12(1-\nu^2)} \mathbf{C}_b \quad (\text{A.3})$$

$$\mathbf{D}_s^{\max} = \frac{\kappa_s Et_{\max}}{2(1+\nu)} \mathbf{I}_2, \quad \mathbf{D}_s^{\min} = \frac{\kappa_s Et_{\min}}{2(1+\nu)} \mathbf{I}_2 \quad (\text{A.4})$$

where $\mathbf{C}_b$ is the normalized bending constitutive matrix, $\mathbf{I}_2$ is the 2×2 identity matrix, and $\kappa_s = 5/6$ is the Reissner shear correction factor. For the physical density $\hat{\rho}_e$, the bending and shear constitutive matrices of each element are interpolated as follows:

$$\mathbf{D}_b^e(\hat{\rho}_e) = \hat{\rho}_e^p \mathbf{D}_b^{\max} + (1 - \hat{\rho}_e^p) \mathbf{D}_b^{\min} \quad (\text{A.5})$$

$$\mathbf{D}_s^e(\hat{\rho}_e) = \hat{\rho}_e^p \mathbf{D}_s^{\max} + (1 - \hat{\rho}_e^p) \mathbf{D}_s^{\min} \quad (\text{A.6})$$

where $p = 3$ is the penalty parameter. This interpolation possesses the following properties:

- $\hat{\rho}_e = 1$ (rib present): $\mathbf{D}_b^e = \mathbf{D}_b^{\max}$, corresponding to the maximum bending stiffness at $t_{\max}$.
- $\hat{\rho}_e = 0$ (base plate only): $\mathbf{D}_b^e = \mathbf{D}_b^{\min}$, corresponding to the baseline bending stiffness at $t_{\min}$.
- Since $t_{\min} > 0$, $\mathbf{D}_b^{\min} > 0$ is guaranteed, preventing the unphysical assumption that plate regions without ribs bear no load whatsoever.

Accordingly, the element stiffness matrix is expressed as follows:

$$\mathbf{K}_e(\hat{\rho}_e) = \int_{\Omega_e} \mathbf{B}_b^T \mathbf{D}_b^e(\hat{\rho}_e) \mathbf{B}_b d\Omega + \int_{\Omega_e} \mathbf{B}_s^{\text{MITC},T} \mathbf{D}_s^e(\hat{\rho}_e) \mathbf{B}_s^{\text{MITC}} d\Omega \quad (\text{A.7})$$

Since both $\mathbf{D}_b^e$ and $\mathbf{D}_s^e$ follow the same interpolation form $\hat{\rho}_e^p(\cdot)^{\max} + (1 - \hat{\rho}_e^p)(\cdot)^{\min}$ with respect to $\hat{\rho}_e$, aggregating at the element stiffness matrix level yields $\mathbf{K}_e(\hat{\rho}_e) = \hat{\rho}_e^p \mathbf{K}_e^{\max} + (1 - \hat{\rho}_e^p) \mathbf{K}_e^{\min}$. Here, $\mathbf{K}_e^{\max}$ and $\mathbf{K}_e^{\min}$ are the element stiffness matrices corresponding to $t_{\max}$ and $t_{\min}$, respectively, and can be precomputed to reduce the stiffness matrix assembly cost at each iteration.

A.3. Density filter

Checkerboard patterns and mesh dependency are well-known numerical pathologies in density-based topology optimization [11]. A checkerboard pattern, in which densities of 0 and 1 alternate between adjacent elements, achieves low compliance mathematically but is impossible to manufacture, constituting an artifact of finite element discretization. To prevent this, a spatially weighted averaging filter is applied to the raw density ρ [44, 45]:

$$\tilde{\rho}_e = \frac{\sum_{j \in N_e} w_{ej} \rho_j}{\sum_{j \in N_e} w_{ej}}, \quad w_{ej} = \max(0, r_{\min} - \text{dist}(e, j)) \quad (\text{A.8})$$

where N_e is the set of neighboring elements within a radius $r_{\min}$ of element e , and w_{ej} is a linear weight inversely proportional to the distance. In this study, $r_{\min} = 5$ (in element size units) is used. This filtering suppresses checkerboard patterns by smoothing abrupt density variations between adjacent elements, and consequently provides indirect control of the minimum feature size to approximately $r_{\min}$. The derivative of the filter is given by:

$$\frac{\partial \tilde{\rho}_e}{\partial \rho_j} = \frac{w_{ej}}{\sum_{k \in N_e} w_{ek}} \quad (\text{A.9})$$

A.4. Maximum feature size control ($r_{\max}$)

While the density filter ($r_{\min}$) controls the minimum member size, restricting the maximum member thickness is also important for manufacturability in rib reinforcement design. Excessively thick ribs cause shrinkage defects during casting and reduce material efficiency. In this study, the local volume constraint approach [62] is adopted to indirectly control the maximum member thickness to approximately $2r_{\max}$. Specifically, a spatial averaging filter with the same structure as the $r_{\min}$ filter is constructed with radius $r_{\max}$, computing the average density in the neighborhood of each element e :

$$\bar{\rho}_e^{\max} = \frac{\sum_{j \in N_e^{\max}} w_{ej}^{\max} \rho_j}{\sum_{j \in N_e^{\max}} w_{ej}^{\max}}, \quad w_{ej}^{\max} = \max(0, r_{\max} - \text{dist}(e, j)) \quad (\text{A.10})$$

If this local average exceeds a threshold $\eta_{\max}$, it indicates that material is overly concentrated in that region. This is suppressed via a quadratic penalty:

$$\mathcal{L}_{\max} = \lambda_{\max} \sum_e [\max(0, \bar{\rho}_e^{\max} - \eta_{\max})]^2 \quad (\text{A.11})$$

where $\lambda_{\max} = 10$ is the penalty weight and $\eta_{\max} = 2V^*$ is the default threshold. The sensitivity of this penalty is efficiently computed through the chain rule of the averaging filter:

$$\frac{\partial \mathcal{L}_{\max}}{\partial \rho_j} = \lambda_{\max} \sum_e \frac{2w_{ej}^{\max}}{\sum_k w_{ek}^{\max}} \max(0, \bar{\rho}_e^{\max} - \eta_{\max}) \quad (\text{A.12})$$

This gradient disperses material away from thick concentration regions, guiding the rib network toward thinner, more distributed patterns.

Because the $r_{\max}$ constraint caps how thick SDS-induced members may become, an overly aggressive setting can force strongly clustered patterns (e.g., thick radial branches) to thin out and converge to a less efficient compromise. The value $r_{\max} = 10$ (with $r_{\min} = 5$) was adopted throughout this study to balance manufacturability against this loss of design freedom.

A.5. Heaviside projection and β -continuation

Density filtering alone does not completely eliminate intermediate densities (gray elements). In fact, the filter tends to smooth density boundaries, widening the gray zone. To address this, the smoothed Heaviside projection [46] is applied to transform the filtered density $\tilde{\rho}_e$ into a physical density $\hat{\rho}_e$ that closely approximates a binarized distribution:

$$\hat{\rho}_e = \frac{\tanh(\beta\eta) + \tanh(\beta(\tilde{\rho}_e - \eta))}{\tanh(\beta\eta) + \tanh(\beta(1 - \eta))} \quad (\text{A.13})$$

where $\eta = 0.5$ is the threshold and β controls the projection sharpness: as $\beta \rightarrow \infty$, the function approaches the ideal Heaviside step function.

Setting β to a large value from the outset causes convergence instability due to severe non-convexity [59]. A continuation strategy is therefore applied in which β starts at $\beta_0 = 1$ and doubles every $N_\beta = 50$ iterations until reaching $\beta_{\max} = 64$:

$$\beta(k) = \min\left(\beta_0 \cdot 2^{\lfloor k/N_\beta \rfloor}, \beta_{\max}\right) \quad (\text{A.14})$$

This β -continuation is particularly important in the present study because the SDS gradient inherently operates toward a continuous image space, tending to induce intermediate densities. A natural two-phase division of labor is established: SDS forms global patterns during the low- β stages, and the subsequent increase in β sharply binarizes these patterns. The effectiveness of this mechanism is quantitatively analyzed in Appendix D.3.

A.6. Compliance sensitivity (physics gradient)

Exploiting the self-adjointness of the compliance $C = \mathbf{F}^\top \mathbf{U}$, the sensitivity with respect to the physical density $\hat{\rho}_e$ is derived as shown in Eq. (A.15), where $\mathbf{u}_e$ and $\mathbf{K}_e$ denote the element-level displacement vector and stiffness matrix, respectively. This sensitivity, further extended via the chain rule in Eq. (A.16), constitutes the physics gradient $\nabla_\rho C$ used in the combined optimization loop.

$$\frac{\partial C}{\partial \hat{\rho}_e} = -\mathbf{u}_e^\top \frac{\partial \mathbf{K}_e}{\partial \hat{\rho}_e} \mathbf{u}_e = -p \hat{\rho}_e^{p-1} \mathbf{u}_e^\top (\mathbf{K}_e^{\max} - \mathbf{K}_e^{\min}) \mathbf{u}_e \quad (\text{A.15})$$

Since $\mathbf{K}_e^{\max} - \mathbf{K}_e^{\min}$ is a constant matrix independent of density, it can be precomputed, and no additional linear system solve is required for the sensitivity computation (self-adjointness of compliance). Extending to the raw design variables ρ_e via the chain rule:

$$\frac{\partial C}{\partial \rho_e} = \frac{\partial C}{\partial \hat{\rho}_e} \cdot \frac{\partial \hat{\rho}_e}{\partial \tilde{\rho}_e} \cdot \frac{\partial \tilde{\rho}_e}{\partial \rho_e} \quad (\text{A.16})$$

where the last two terms represent the chain rule contributions of the Heaviside projection (Appendix A.5) and the density filter (Appendix A.3), respectively. This complete sensitivity is denoted by the physics gradient $\nabla_\rho C$ in this study and constitutes the key component combined with the SDS gradient. The quantity $\mathbf{u}_e^\top (\mathbf{K}_e^{\max} - \mathbf{K}_e^{\min}) \mathbf{u}_e$ represents the strain energy associated with the stiffness gain from adding a rib to element e . Elements with larger values of this quantity yield greater compliance reduction upon rib placement, and thus the sensitivity provides information on where rib placement would be most efficient.

A.7. Finite-difference verification of the thermoelastic sensitivity

For the thermoelastic objective the load $\mathbf{F}_{\text{th}}$ is assembled from the prescribed moment $\mathbf{M}_{\text{th}} = \alpha_T \Delta T [1, 1, 0]^\top$ without any dependence on the element stiffness $E(\hat{\rho})$. Consequently $\partial \mathbf{F}_{\text{th}} / \partial \hat{\rho} = \mathbf{0}$: the load carries no design-derivative term, and the self-adjoint sensitivity of Eq. (A.15) applies unchanged, with only the stiffness $\mathbf{K}(\hat{\rho})$ being density-dependent. To confirm this numerically, the analytic sensitivity was compared against a central finite difference $[C(\hat{\rho}_e + h) - C(\hat{\rho}_e - h)] / (2h)$ with $h = 10^{-4}$ at 30 randomly selected elements of a mid-density field ($\hat{\rho}_e \sim \mathcal{U}[0.3, 0.7]$). The analytic and finite-difference gradients show excellent agreement (maximum relative error 3.2×10^{-4} , mean 7.0×10^{-5} , at the level of the central-difference truncation error; Pearson correlation 0.99999995). In the adopted prescribed-thermal-moment formulation the assembled thermal load is density-independent, so the standard self-adjoint sensitivity is consistent with this formulation and no load-derivative correction is required.

B. SDS coupling and optimization details

B.1. Density field to image conversion

Section 3.3 describes the conversion of $\hat{\rho} \in \mathbb{R}^{n_{\text{elx}} \times n_{\text{ely}}}$ into the diffusion-model latent $\mathbf{z}_0 = \mathcal{E}(\mathbf{x}) \in \mathbb{R}^{4 \times 64 \times 64}$ via three-channel replication, bilinear resize to 512×512 , and VAE encoding. This pipeline is fully

differentiable. The chain rule for backpropagation of the SDS gradient from $\mathbf{z}_0$ to $\hat{\rho}$ is

$$\frac{\partial \mathbf{z}_0}{\partial \hat{\rho}} = \frac{\partial \mathcal{E}}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial \hat{\rho}} \quad (\text{B.1})$$

The full SDS pipeline, including the forward diffusion, classifier-free guidance, and back-propagation path, is illustrated in Fig. 2.

B.2. EMA smoothing

The SDS gradient is inherently stochastic. Since the timestep t and noise ϵ are randomly sampled at each iteration, considerable variance exists in both the direction and magnitude of the gradient. In particular, under the setting of this study where a single-sample gradient is used without mini-batch averaging, this variance can undermine optimization stability. To mitigate this, EMA smoothing is applied as defined in Eq. (9) (Section 3.4). The smoothing coefficient is set to $\alpha = 0.9$, with gradients exponentially averaged over approximately the most recent 10 iterations. This provides an appropriate balance between smoothing the noise of individual iterations and reflecting the temporal evolution of the SDS gradient due to timestep annealing. The ablation study (Section 6.5) shows that $\alpha = 0$ (no EMA) results in +5.0% degradation and $\alpha = 0.5$ yields +3.4% degradation.

B.3. Volume constraint enforcement

A bisection projection is employed to satisfy the volume fraction constraint $V(\hat{\rho}) = V^*$ exactly at each iteration. After the *Adam* optimizer updates the densities, a uniform shift Δ is applied to all design variables such that the projected volume matches the target:

$$\rho_e^{(k+1)} \leftarrow \rho_e^{(k+1)} + \Delta, \quad \text{where } \Delta \text{ solves } \frac{1}{N_{\text{active}}} \sum_{e \in \Omega_{\text{active}}} \hat{\rho}_e(\rho^{(k+1)} + \Delta) = V^* \quad (\text{B.2})$$

While traditional topology optimization uses the Lagrangian multiplier method [3] or *MMA* [57] to handle volume constraints, this direct bisection approach is adopted, given the compatibility with the *Adam* optimizer and the stochastic nature of the SDS gradient. The scalar Δ is found via bisection on the monotone left-hand side of Eq. (B.2). Here, Ω_{active} denotes the set of active elements within the domain (excluding elements outside the circle in D_{Circle}). The bisection exploits the monotonicity of the Heaviside projection and converges to machine precision within approximately 30 iterations. This method fixes the volume fraction exactly at the target value, eliminating ambiguity caused by volume-fraction differences when comparing compliance.

B.4. Optimizer: *Adam* vs. *OC*

The *Adam* optimizer [54] is used in this study instead of *OC*, which is widely employed in conventional topology optimization. There are two reasons for this choice. First, it offers good compatibility with the SDS gradient. *OC* is a heuristic update rule that exploits the special structure of compliance sensitivity (self-adjointness and non-negativity), making it difficult to integrate arbitrary gradients, such as the SDS gradient, in a natural way. *Adam* tracks the first and second moments of arbitrary gradients and applies an adaptive learning rate, making it well suited for combining physics and SDS gradients. Second, it converges stably in conjunction with the Heaviside projection. In experiments, *OC* combined with Heaviside was observed to exhibit period-2 oscillation at high β values without converging (see Fig. B.1). This occurs because the bisection update of *OC* interacts with the sharp nonlinearity of the Heaviside projection, causing oscillation between two stable attractors. *Adam* naturally damps this oscillation through momentum and learning rate decay ($0.04 \rightarrow 0.01$, cosine annealing). The momentum coefficients ($\beta_1 = 0.5$, $\beta_2 = 0.9$) are set lower than the standard defaults (0.9, 0.999) to shorten the gradient memory, allowing the optimizer to adapt quickly as the SDS weight and timestep change during the scheduling phases. This behavior is illustrated in Fig. B.1.

The same SIMP optimization (without SDS) was conducted with the conventional *OC* optimizer across all three plate-bending domains. The quantitative comparison appears in Table B.1 and the corresponding density fields in Fig. B.2. The most severe case is D_{Circle} , where *OC* compliance oscillates between $C = 10.639$ and $C = 8.393$ in the final iterations, and the volume fraction collapses to $V = 0.13$ (target $V^* = 0.20$),

Zero-shot rib design

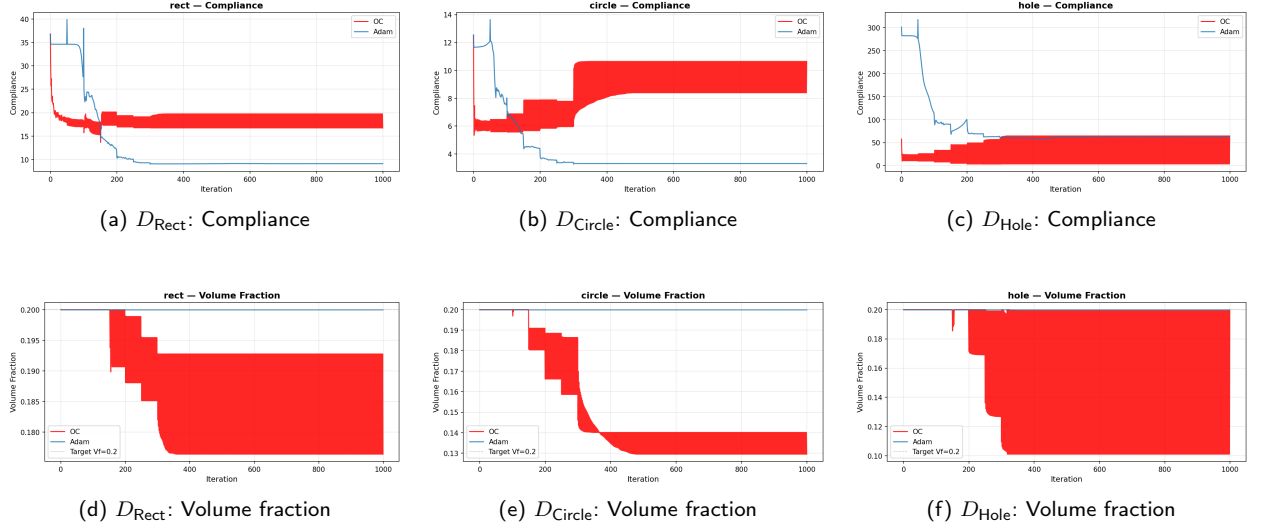


Figure B.1: Convergence comparison of *OC* (oscillatory) vs. *Adam* (stable) under Heaviside projection.

resulting in a nearly empty density field (Fig. B.2b). In D_{Hole} , *OC* compliance ($C = 148.959$) is $2.36\times$ worse than the *Adam* baseline ($C = 63.113$), with severe fragmentation visible in the density field (Fig. B.2c). In D_{Rect} , *OC* produces a recognizable but disconnected structure with compliance $1.84\times$ worse than *Adam* (Fig. B.2a). The *Adam* baseline therefore represents a strong conventional SIMP result, justifying the choice of *Adam* over *OC* for the Heaviside-projected SIMP framework used throughout this study.

Table B.1

Baseline compliance comparison: *Adam* vs. *OC* optimizer (no SDS, $V^* = 0.20$).

Domain	<i>Adam</i> C	<i>OC</i> C	<i>OC</i> / <i>Adam</i>
D_{Rect}	9.106	16.736	$1.84\times$
D_{Circle}	3.308	10.639	$3.22\times$
D_{Hole}	63.113	148.959	$2.36\times$

C. Post-processing details: density field to CAD geometry

The optimized density field $\hat{\rho}$ is a continuous 2D image that cannot be directly manufactured. A six-stage pipeline is presented to automatically convert it into a CAD-convertible candidate geometry (STEP format). The complete pipeline is illustrated in Fig. C.1. This optimization-to-CAD/CAE (computer-aided engineering) integration shares a similar objective to prior work [63] that automated a 3D conceptual wheel design-evaluation pipeline in deep-learning-based generative design. However, the present study adopts a skeleton-based approach tailored to density-field topology-optimization results. The entire pipeline completes in 5 seconds on a single CPU core and requires no user parameters beyond the extrusion height.

Step 1: Binarization The continuous density field is converted into a binary mask M using a threshold of $\rho_{\text{th}} = 0.5$:

$$M(i, j) = \begin{cases} 1 & \text{if } \hat{\rho}(i, j) > 0.5 \\ 0 & \text{otherwise} \end{cases}$$

Since the Heaviside projection ($\beta = 64$) already produces a nearly binary distribution (gray density $< 3\%$, as shown in Appendix D.3), additional morphological post-processing (closing, opening) is unnecessary.

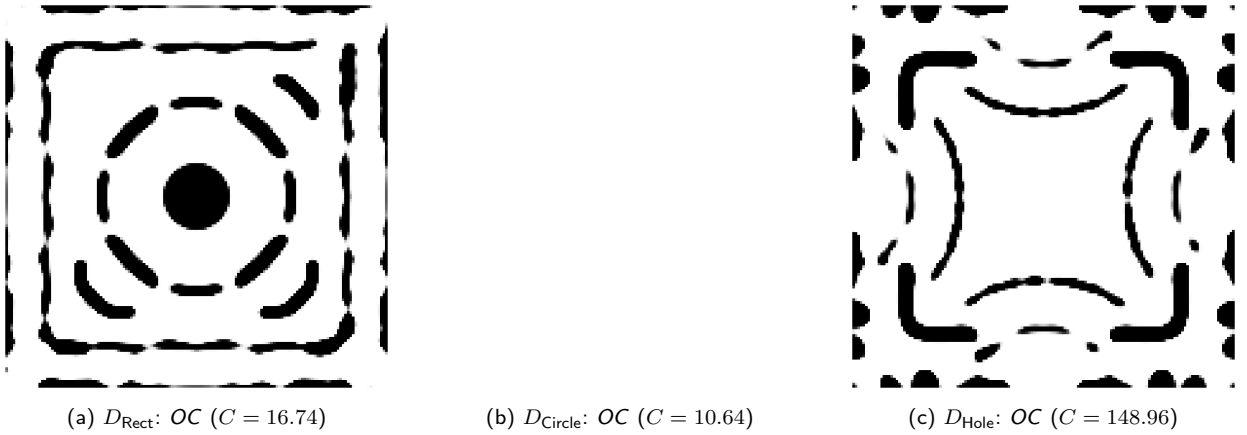


Figure B.2: OC baseline density fields (no SDS). (a) D_{Rect} : disconnected rib pattern with compliance $1.84\times$ worse than *Adam*. (b) D_{Circle} : catastrophic failure due to period-2 oscillation; volume fraction collapses to $V = 0.13$, yielding a nearly empty domain. (c) D_{Hole} : severely fragmented density field with compliance $2.36\times$ worse than *Adam*.

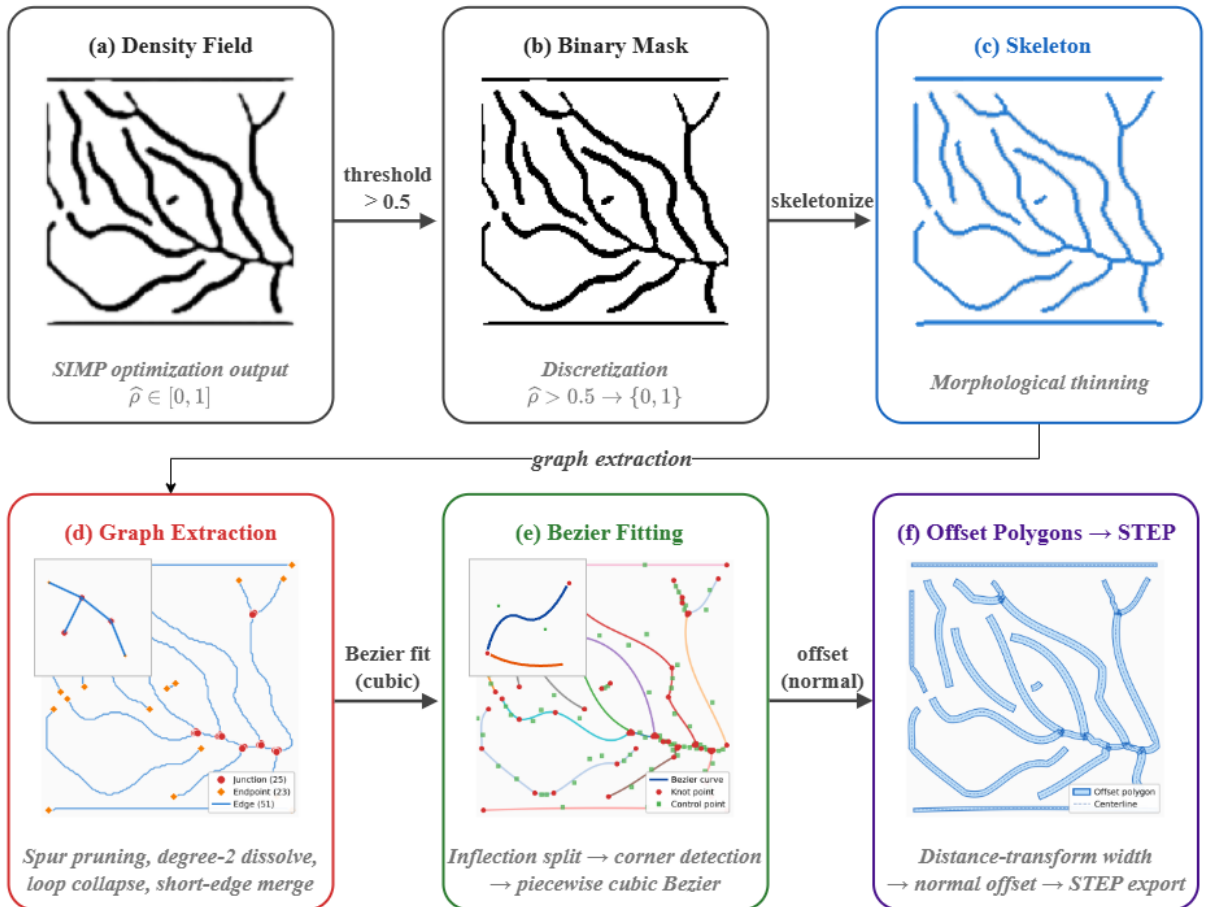


Figure C.1: Post-processing pipeline: density field $\rightarrow$ binarization $\rightarrow$ skeletonization $\rightarrow$ graph $\rightarrow$ Bézier fitting $\rightarrow$ CAD export.

Step 2: Skeletonization and Distance Transform Two operations are performed in parallel on the binary mask. First, Zhang-Suen thinning [55] extracts the one-pixel-wide medial axis (skeleton) of the foreground region. This skeleton captures the topological structure of the rib network, including branching patterns and connectivity. Second, a Euclidean distance transform computes the distance from each foreground pixel to the nearest background pixel. The distance value at skeleton pixels corresponds to the local half-thickness of the respective rib, which is used in *Step 5* to assign physically meaningful rib widths.

Step 3: Graph Extraction The skeleton image is converted into a topological graph $G = (V, E)$. Each skeleton pixel is classified by its degree in the 8-connected neighborhood: pixels with degree ≥ 3 are registered as branch point nodes, and pixels with degree $= 1$ as endpoint nodes. Skeleton paths between nodes are traced via depth-first search to construct edges E , where each edge is represented as $(v_a, v_b, \mathbf{P})$ with $\mathbf{P} \in \mathbb{R}^{n_e \times 2}$ being the pixel coordinate sequence of the skeleton path.

Step 4: Bézier Curve Fitting Each edge path $\mathbf{P}$ is approximated by a piecewise cubic Bézier curve. First, points where the local direction changes by more than 30° within an adaptive window are detected as corners, thereby splitting the path into sub-segments. For each segment, the two endpoints are fixed, and the internal control points $\mathbf{P}_1, \mathbf{P}_2$ are determined via least-squares fitting based on chord-length parameterization.

$$\mathbf{B}(\xi) = (1 - \xi)^3 \mathbf{P}_0 + 3(1 - \xi)^2 \xi \mathbf{P}_1 + 3(1 - \xi) \xi^2 \mathbf{P}_2 + \xi^3 \mathbf{P}_3$$

All control points are converted from pixel coordinates to physical domain coordinates (x, y) .

Step 5: Offset Polygon Generation The Bézier centerline of each edge is expanded into a closed 2D polygon with physically meaningful thickness. The distance transform values are sampled along the skeleton path of each edge, and their mean is taken as the half-width w_{half} of the edge (in pixel units). This is converted to domain coordinates using the average pixel-to-domain scale factor $s = (W/n_{\text{elx}} + H/n_{\text{ely}})/2$. A minimum half-width of 0.3 is enforced to prevent degenerate polygons. Each Bézier segment is evaluated at 20 equally spaced parameter values to obtain centerline points and tangent vectors. At each point, the unit normal $\mathbf{n} = (-\dot{y}, \dot{x})/|\dot{\mathbf{B}}|$ is computed, and left/right offset curves are generated:

$$\mathbf{C}_{\text{left}} = \mathbf{B}(\xi) + w_{\text{half}} \cdot \mathbf{n}(\xi), \quad \mathbf{C}_{\text{right}} = \mathbf{B}(\xi) - w_{\text{half}} \cdot \mathbf{n}(\xi)$$

The left curve and the reversed right curve are connected to form a closed polygon.

Step 6: CAD Export The 2D polygons are extruded into 3D solids to ensure CAD compatibility. Each polygon is converted into a closed wire, a planar face is constructed, and the face is extruded by a specified height h (default 5 mm) in the z -direction. All solids are exported as an AP214 schema (ISO 10303-214) STEP data file. Optionally, a base plate with a configurable thickness can be added beneath the rib structure.

D. Additional results

D.1. Alternative baselines: numerical summary

Best-of-run compliance values underpinning the comparison in Section 5.2 appear in Table D.1.

A best-of-run comparison alone could be confounded by the number of restarts allocated to each method. To control for this, the two SIMP-only baselines were re-run under a matched budget of 50 random restarts (multi-start) and 50 periodic-perturbation runs ($\sigma = 0.05$) per domain, then compared against the five SDS seeds by resampling equal-sized five-run subsets. Fig. D.1 shows the resulting distributions. In none of 5000 resampled five-run subsets, for either baseline in any of the three domains, does the best baseline compliance reach the best SDS compliance. Indeed, even the best of all 50 runs per baseline falls short of the best SDS result in every domain, so no resampled five-run subset reaches it (0 of 5000). The baseline distributions lie at or above the SDS result, and in D_{Circle} both baselines remain near the SIMP baseline compliance. The SDS improvement is therefore not reproducible by additional unstructured search under a matched number-of-runs budget. Wall-clock costs are reported separately because SDS and SIMP use different CPU/GPU resources.

Table D.1

Best-of-run compliance C for the proposed SDS framework against three SIMP-only alternatives (lower is better). All methods share the same SIMP solver and therefore require no additional training. The multi-start and perturbation entries are the best of 50 runs per domain; a matched-budget distributional comparison is given in Fig. D.1. [†]Best SDS run on D_{Hole} is at $\lambda_{\text{sds}}^0 = 50$ (see Appendix D.2); the remaining domains use $\lambda_{\text{sds}}^0 = 10$.

Method	D_{Rect}	D_{Circle}	D_{Hole}
Baseline SIMP (<i>Adam</i>)	9.106	3.308	63.113
Multi-start SIMP (best)	8.102	3.166	47.434
Perturbation SIMP (best)	8.339	3.117	47.495
Image-guided (best)	9.217	3.105	49.505
SDS (best)	7.886	2.522	43.20[†]

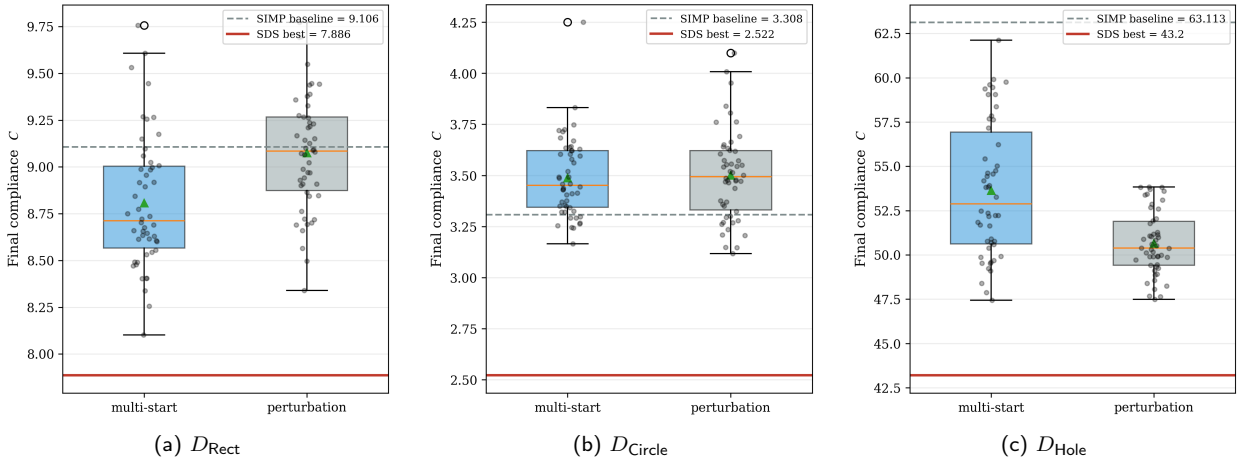


Figure D.1: Matched-number-of-runs comparison of the SIMP-only baselines against SDS across the three plate-bending domains. Box plots show the final compliance of 50 multi-start and 50 perturbation runs per domain (individual runs jittered; green triangle = mean); the dashed line is the SIMP baseline and the solid red line the best SDS compliance. In every domain the best SDS result lies below the entire baseline distribution, so neither multi-start nor perturbation reaches it under a matched budget.

For full transparency of the per-seed results underlying the mean $\pm$ std values in Table 5, Figs. D.2, D.3, and D.4 show the compliance of every individual seed (five per prompt–domain combination) as strip-and-box plots, with the SIMP-only baseline marked. Prompt and hyperparameter selection was carried out on separate single-seed pilot runs before this five-seed statistical evaluation, keeping model selection distinct from final testing.

D.2. D_{Hole} SDS weight ($\lambda_{\text{sds}}^0 = 50$)

D_{Hole} uses $\lambda_{\text{sds}}^0 = 50$ rather than the $\lambda_{\text{sds}}^0 = 10$ adopted for D_{Rect} , D_{Circle} , and D_{Link} . Because the baseline compliance ($C = 63.113$) is roughly seven times that of D_{Rect} , the physics-gradient magnitude is correspondingly larger. Under the raw-weighted-sum update, λ_{sds}^0 must rise to keep the generative signal competitive. The scaling rationale is identical to the thermoelastic case (Appendix D.6). Full per-prompt results at $\lambda_{\text{sds}}^0 = 50$ appear in Table 5. Domain-level interpretation is developed in Section 5.1.

D.3. Gray density: quantitative ablation

This section reports the quantitative ablation of Heaviside projection and β -continuation. The mechanistic origin of the SDS-induced gray-density tendency is discussed in Section 6.4.

Without Heaviside projection, SDS guidance yields a gray-density ratio of 42.6% against 20–21% for pure SIMP, nearly doubling the baseline. β doubles every 50 iterations following $\beta_k = \min(2^{\lfloor k/50 \rfloor}, 64)$. Fig. 13

Zero-shot rib design

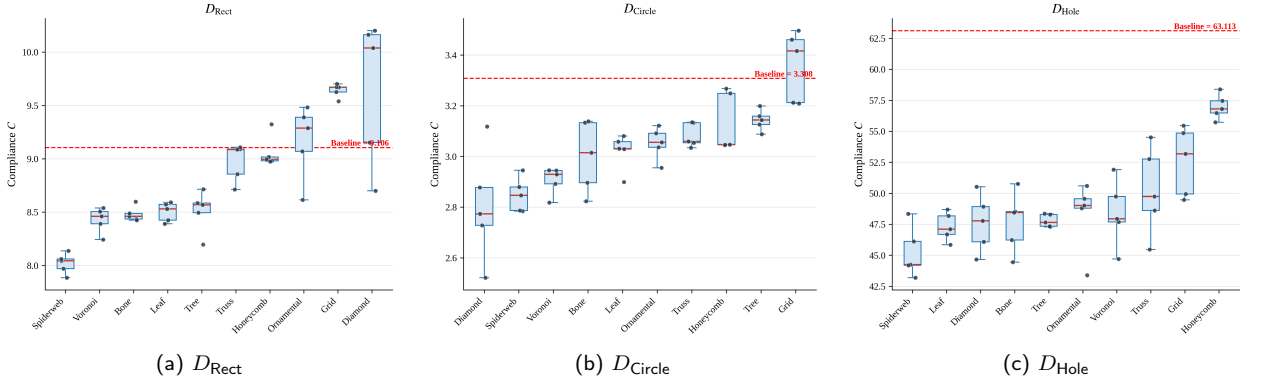


Figure D.2: Seed-level compliance for the three mechanical plate-bending domains, all ten prompts, five seeds each (dots: individual seeds; box: interquartile range; red line: median). Prompts are ordered by mean compliance; the dashed line is the SIMP-only baseline.

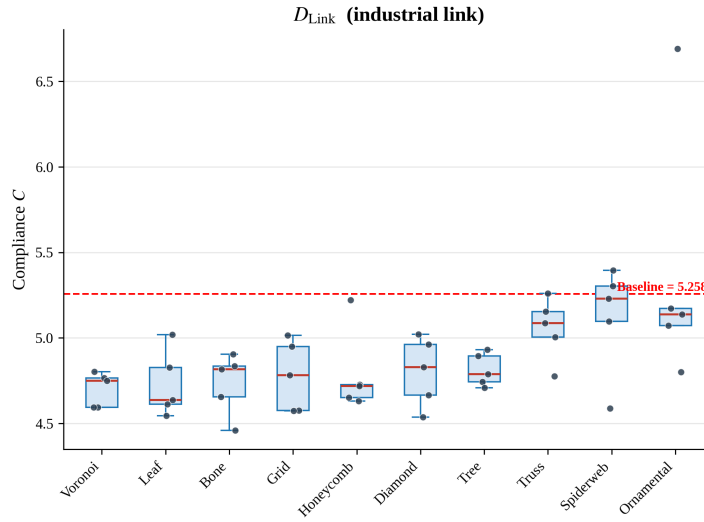


Figure D.3: Seed-level compliance for the industrial D_{Link} domain (ten prompts, five seeds each; baseline $C = 5.258$ dashed).

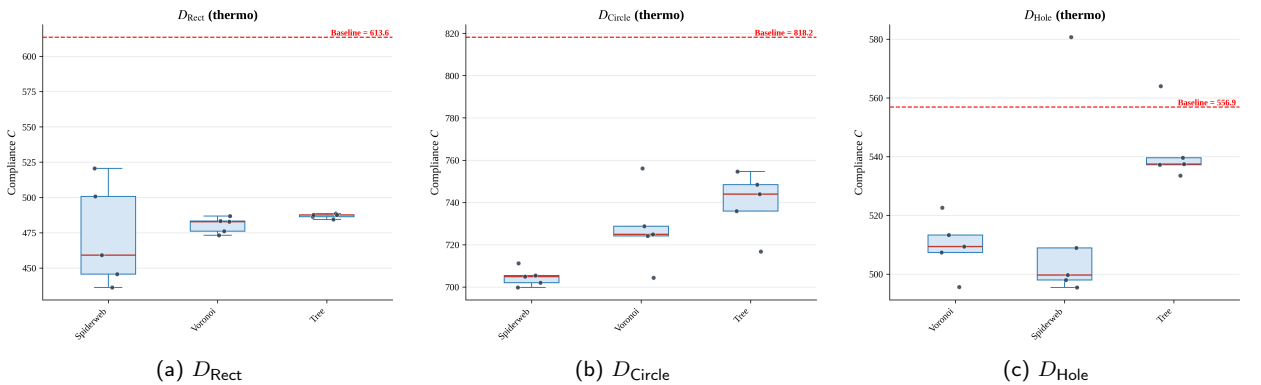


Figure D.4: Seed-level thermoelastic compliance for the three plate-bending domains (three prompts each, five seeds); every prompt improves on the baseline (dashed).

and Table D.2 report the gray-density/compliance trade-off for $\beta_{\max} \in \{32, 64\}$ against the projection-off baseline. Higher $\beta_{\max}$ drives the gray ratio below 3% with no compliance penalty (Table D.2: $C = 9.29$ at $\beta_{\max} = 64$ versus 9.93 at $\beta_{\max} = 32$), which justifies the choice $\beta_{\max} = 64$.

Table D.2

Effect of Heaviside projection and β -continuation on compliance and gray density.

Experiment	Heaviside	$\beta_{\max}$	Compliance	Gray ratio (%)
A	OFF	—	18.66	42.6
B	ON	64	9.29	2.5
C	ON	32	9.93	5.1

D.4. When SDS fails: over-constraining

Conversely, there are cases where SDS degrades performance. Two representative failure modes are observed in D_{Rect} . The *Diamond* prompt (+6.0%) over-constrains the density field into a rigid rhombus lattice that conflicts with the natural diagonal load paths from center to corners. Similarly, the *Grid* prompt (+5.9%) forces an orthogonal lattice that wastes material in regions of low stress concentration. In both cases, the SDS gradient consistently opposes the physics gradient: while the physics gradient seeks to concentrate material along load paths, the SDS gradient redistributes it to maintain the prompted pattern, leading to a compromise that is optimal for neither.

The cosine cooldown mitigates this conflict by gradually reducing SDS influence. However, without cooldown, such conflicts persist until the final convergence, as confirmed in the scheduling ablation (Section 6.5): removing the timestep annealing alone causes +13.0% degradation even for the well-aligned *Spiderweb* prompt. For misaligned prompts, the degradation would be more severe. These observations suggest a practical guideline: when the prompted pattern is uncertain in its compatibility with the load paths, a shorter cooldown or lower λ_{sds} can serve as a conservative safeguard.

D.5. Prompt ablation: isolating the semantic contribution

The contribution of the prompt’s *semantic content* is isolated by varying the prompt while holding the solver and the SDS coupling fixed. On D_{Rect} , the five control conditions share the same five random seeds ($\lambda_{\text{sds}}^0 = 10$), so they are compared under identical initialization and noise realizations (a matched-seed comparison); the *Spiderweb* reference is the main-experiment value (−11.9%, Table 5). The analysis relies only on the relative ranking across the six conditions. The best structural prompt (*Spiderweb*) is placed alongside five controls: an *empty* prompt (unconditional SDS), a *prefix-only* prompt (“bold black lines on white background,” with the structural noun removed), two semantically irrelevant prompts (*fruit*, *sunset*), and a *dot* prompt that changes a single word (“lines”→“dots”) so that the prior forms isolated points rather than connected line networks. Table D.3 reports the final compliance, win rate, and mean skeleton endpoint count n_{end} . The corresponding rib morphologies and the n_{end} –compliance correlation appear in Fig. 10 (Section 6.1).

Three findings emerge. First, the structural prompt (−11.9%) far outperforms the empty and prefix-only controls (−3.6% and −4.5%), and the improvement tracks the endpoint count that structural prompts minimize rather than the mere presence of a generative term. The semantically irrelevant *fruit* prompt still helps (−5.5%) by yielding a relatively low-endpoint morphology. Second, the *dot* prompt degrades compliance the most (+7.3%, 0/5 seeds winning), raising the endpoint count to the highest of any condition ($n_{\text{end}} = 49$ versus 19 for *Spiderweb*). The semantically irrelevant *sunset* prompt, with the next-highest endpoint count ($n_{\text{end}} = 36$), is the only other condition to worsen the baseline (+3.6%), so the two most endpoint-rich prompts are precisely the two that degrade compliance. Third, across the thirty individual runs (five seeds per prompt) the endpoint count and the compliance change are strongly correlated (Pearson $r = 0.87$). This controlled variation is consistent with a causal role for the endpoint count in the dead-end suppression mechanism of Section 6.1. The within-design surgery of Appendix D.9, which manipulates the endpoints without altering the surrounding morphology, provides the direct intervention-based support.

Table D.3

Prompt ablation on D_{Rect} (mean $\pm$ std over five seeds; baseline $C = 9.106$). Conditions are ordered by mean skeleton endpoint count n_{end} . “Wins” counts seeds improving over the baseline. The two conditions with the highest n_{end} (*dot*, *sunset*) degrade compliance, the *dot* prompt most.

Prompt condition	C	$\Delta\%$	Wins	n_{end}
Semantic (<i>Spiderweb</i>)	8.02 ± 0.10	-11.9%	5/5	19.2
Random (<i>fruit</i>)	8.61 ± 0.31	-5.5%	5/5	22.6
Empty (unconditional)	8.77 ± 0.23	-3.6%	4/5	25.8
Prefix-only	8.70 ± 0.29	-4.5%	5/5	26.4
Random (<i>sunset</i>)	9.43 ± 0.36	$+3.6\%$	2/5	36.2
Dot (isolated points)	9.77 ± 0.34	$+7.3\%$	0/5	49.2

D.6. λ_{sds}^0 sensitivity in thermoelastic loading

To empirically locate the optimal SDS weight for thermoelastic loading, the D_{Rect} domain was swept across $\lambda_{\text{sds}}^0 \in \{10, 50, 100, 200\}$ using the *Spiderweb* prompt as a consistent probe (5 seeds per cell). The value adopted for all three prompts in the thermoelastic experiments ($\lambda_{\text{sds}}^0 = 100$) was selected based on this single-prompt sweep. The results appear in the lower block of Table 10: a clear U-shape emerges, in which small λ_{sds}^0 provides insufficient generative guidance against the larger and more diffuse thermal physics gradient, the improvement grows monotonically up to $\lambda_{\text{sds}}^0 \approx 100$ (best -23.0%), and $\lambda_{\text{sds}}^0 = 200$ already shows mild saturation (-20.0%), consistent with excessive SDS overriding the physics gradient. This U-shape mirrors the λ_{sds} ablation observed in the mechanical setting (upper block of the same table) but is shifted roughly tenfold toward larger values, quantifying the claim that the optimal SDS weight scales with the magnitude of the physics gradient under the raw weighted sum (Section 5.5).

D.7. Mesh refinement and the filter length scale

Section 3.1 notes that fixing the filter radius in physical units reduces mesh-dependent features. To examine this beyond the single 150×150 resolution used in the main experiments, a mesh-refinement study was carried out on D_{Hole} at three resolutions, $N \in \{100, 150, 200\}$. The essential requirement is that the filter radius be fixed in *physical* units rather than in element counts: setting $r_{\text{min}} = N/30$ and $r_{\text{max}} = N/15$ holds the minimum feature size constant under refinement, whereas a fixed element-count radius would shrink the physical feature size and reintroduce mesh-dependent fine detail. D_{Hole} is used here because its four-sided distributed loading provides a well-conditioned setting for assessing mesh convergence.

Table D.4 and Fig. D.5 summarize the outcome, and two observations provide preliminary evidence of mesh-refinement robustness. First, the baseline SIMP topology is qualitatively consistent across resolutions (Fig. D.5, first column): the same four curved ribs framing the central hole emerge at both resolutions, confirming that the physical-unit filter fixes the structural length scale. Second, the SDS improvement over the same-resolution baseline persists at every resolution (for the *Spiderweb* probe, -19.3% , -28.4% , and -10.0% at $N = 100, 150, 200$, mean over 5 seeds), so the improvement is not confined to a single discretization.

At each resolution, λ_{sds}^0 was scaled with the baseline compliance following the convention adopted across domains (Section 3.3). Because the density field is resized to 512×512 before entering the diffusion model, the SDS gradient magnitude is essentially resolution-independent, whereas the physics-gradient magnitude grows with the baseline compliance. Scaling λ_{sds}^0 accordingly keeps the two gradients in a comparable balance as the mesh is refined.

D.8. 3D re-analysis of the protruding-rib geometry

The optimization models a reinforced region as a locally stiffened Mindlin plate, whereas the exported geometry is a base plate with *protruding* ribs whose bending stiffness derives from a neutral-axis offset. To verify that this 2D model preserves the design comparisons under a 3D protruding-rib solid representation reconstructed from the converged density field, each field was re-analyzed as a 3D solid: it was extruded into a base plate (thickness 1) carrying ribs (height 5) on the material elements and solved with 3D linear elasticity in FEniCSx under the same loading and support as the corresponding 2D problem for each domain.

Table D.4

Mesh refinement of D_{Hole} with the filter radius fixed in physical units ($r_{\min} = N/30$, $r_{\max} = N/15$). C_{base} and C_{SDS} (mean $\pm$ std over 5 seeds) are the SIMP-only and SDS-guided compliances; $\Delta\%$ is the improvement over the same-resolution baseline. The *Spiderweb* prompt is used as a consistent probe.

N	$r_{\min}$	C_{base}	λ_{SDS}^0	C_{SDS}	$\Delta\%$
100	3.33	26.65	22.7	21.50 ± 1.45	-19.3%
150	5.00	63.11	50.0	45.22 ± 2.04	-28.4%
200	6.67	86.69	73.9	78.01 ± 3.67	-10.0%

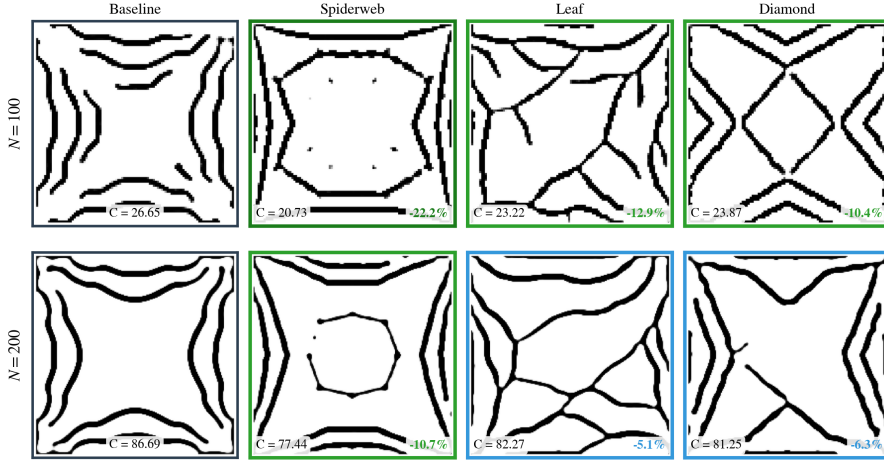


Figure D.5: Mesh refinement of D_{Hole} with the filter radius fixed in physical units ($r_{\min} = N/30$). Rows: $N = 100$ and $N = 200$ ($N = 150$ appears in Fig. 3 and Table 5). Columns: the same-resolution SIMP baseline and three representative prompts (*Spiderweb*, *Leaf*, *Diamond*; median seed). Cell-border color encodes the improvement over the baseline; per-panel compliance C and $\Delta\%$ are overlaid.

Table D.5 and Fig. D.6 report the outcome across the three plate-bending domains. The compliance predicted by the 2D plate model correlates positively with that of the full 3D solid re-analysis in every domain (Pearson $r = 0.92, 0.82$, and 0.98 for D_{Rect} , D_{Circle} , and D_{Hole}), and the top-ranked prompt coincides between the two models in D_{Rect} (*Spiderweb*). In D_{Circle} and D_{Hole} the best 2D prompt remains among the strongest designs in 3D. The relative ranking of prompts is thus largely preserved under the 3D solid representation, indicating that the 2D plate model retains substantial predictive value for comparative ranking, although this does not by itself eliminate the modeling discrepancy between locally thickened plates and protruding-rib solids. A full density–binary–STEP comparison including mass, volume fraction, and deformation fields would provide more comprehensive validation and remains future work.

Table D.5

Agreement between the 2D plate model and the full 3D solid re-analysis of the protruding-rib geometry, across the three plate-bending domains. C_{base} is the SIMP-only baseline compliance (2D). Correlation is Pearson r between the compliance values predicted by the two models over the prompt set (baseline included).

Domain	C_{base} (2D)	n prompts	Pearson r
D_{Rect}	9.106	10	0.92
D_{Circle}	3.308	10	0.82
D_{Hole}	63.113	10	0.98

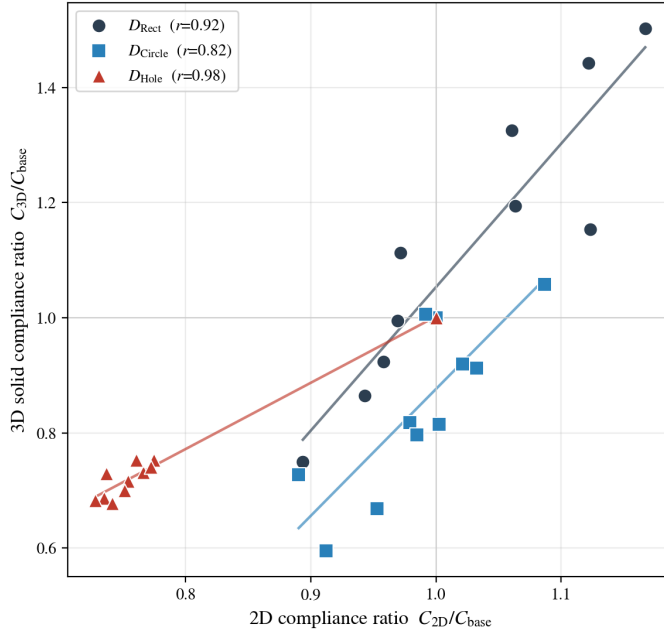


Figure D.6: Compliance of the 2D plate model (horizontal) versus the full 3D solid re-analysis of the protruding-rib geometry (vertical), both normalized by the SIMP-only baseline compliance, for the three plate-bending domains under matched loading and support. Points below and to the left of the baseline (ratio = 1) improve over the baseline in both models; the positive correlation and per-domain trend lines show that the 2D design ranking is largely preserved in 3D.

D.9. Within-design causal test of dead-end removal

The prompt ablation (Appendix D.5) shows that prompts inducing more skeleton endpoints yield worse compliance, and the *dot* prompt establishes this by natural variation. A differentiable endpoint-count penalty is not available in the present implementation because the skeletonization that defines endpoints is non-differentiable, so a direct post-optimization intervention is used instead to isolate the mechanical role of endpoint branches. To rule out any confound from the prompt also reshaping the global morphology, on a fixed optimized design only the skeleton endpoints are altered and re-analyzed by FEA. Three interventions are applied to nine designs (three seeds in each of D_{Rect} , D_{Circle} , and D_{Hole}): (i) *prune dead-ends*, in which every skeleton-endpoint twig, traced to its nearest branch point, is removed together with its local thickness; (ii) *prune a load-bearing rib*, a control in which the single non-dead-end rib carrying the most strain energy is removed. And (iii) *graft a dead-end*, in which an artificial stub is added at a branch point toward the void. Compliance is evaluated with the same plate FEA and boundary conditions used throughout.

Fig. D.7 illustrates the interventions, and the per-design means are collected in Table 8 (Section 6.1). The contrast is consistent across all nine designs. Pruning dead ends raises compliance by only +5.9% on average (range +0.6 to +12.1%), whereas pruning a single load-bearing rib raises it by +61.5% on average (up to +182% in D_{Hole}). Grafting an artificial dead-end changes compliance by less than 1% (−0.9 to −0.1%), directly showing that a dead-end adds no meaningful load path. Consistently, the pruned dead-end regions carry only 0.07–3.7% of the total strain energy, versus 8.3–49.5% for the removed load-bearing rib. The small but non-zero cost of dead-end removal is the expected first-order effect ($\Delta C \approx p f$ when removing material of strain-energy fraction f at penalization $p = 3$). Because dead ends carry only a few percent of the strain energy, removing them costs roughly an order of magnitude less than removing a true rib. This intervention leaves the surrounding topology unchanged and thus confirms causally, without the morphological confound of the *dot* prompt, that dead-end suppression underlies the SDS-guided improvement.

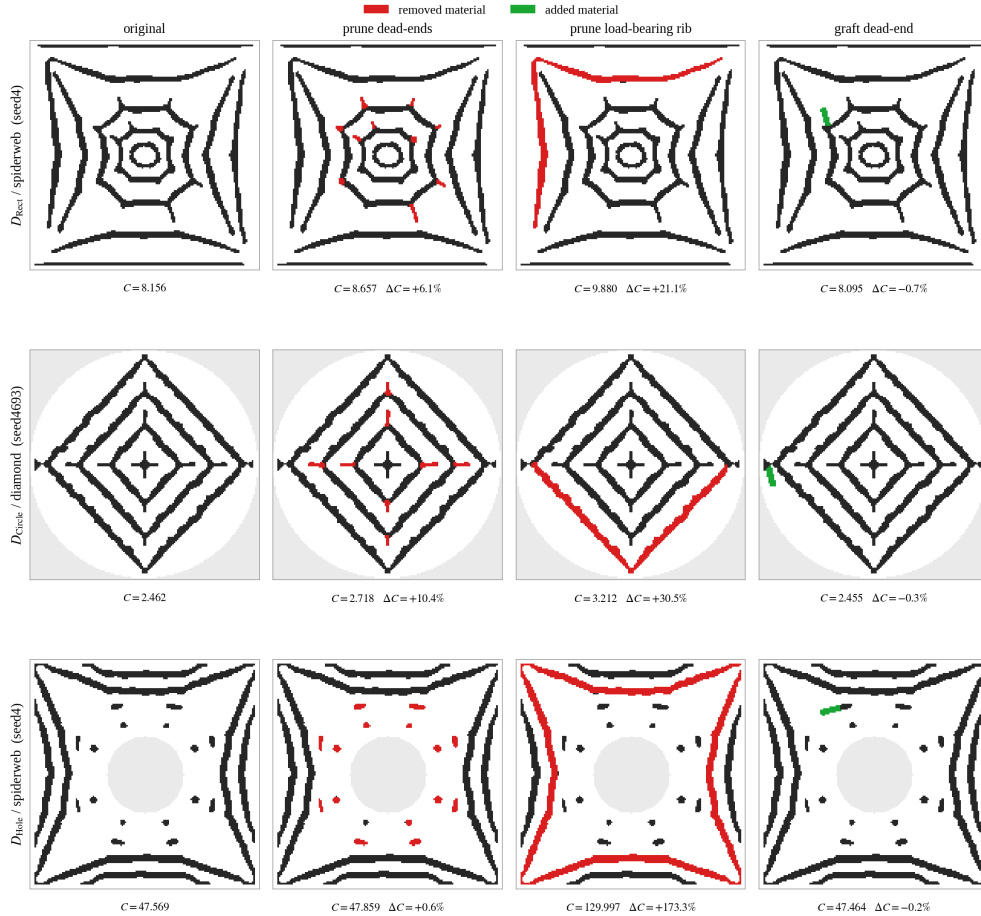


Figure D.7: Within-design causal test of dead-end (skeleton-endpoint) removal, one representative design per domain. Columns: original; dead-ends pruned (removed material in red); a single load-bearing rib pruned (red); an artificial dead-end grafted (added material in green). Each panel is labeled with its compliance C and change ΔC . Removing dead ends barely changes C and grafting one is inert, whereas removing a load-bearing rib is catastrophic (up to +182% in D_{Hole}), establishing that dead ends carry comparatively little load.

References

- [1] M. P. Bendsoe, O. Sigmund, *Topology optimization: theory, methods, and applications*, Springer Science & Business Media, Berlin, Heidelberg, 2013.
- [2] M. P. Bendsoe, O. Sigmund, Material interpolation schemes in topology optimization, *Archive of applied mechanics* 69 (9) (1999) 635–654.
- [3] M. P. Bendsoe, Optimal shape design as a material distribution problem, *Structural optimization* 1 (4) (1989) 193–202.
- [4] O. Sigmund, A 99 line topology optimization code written in matlab, *Structural and multidisciplinary optimization* 21 (2) (2001) 120–127.
- [5] E. Andreassen, A. Clausen, M. Schevenels, B. S. Lazarov, O. Sigmund, Efficient topology optimization in matlab using 88 lines of code, *Structural and Multidisciplinary Optimization* 43 (1) (2011) 1–16.
- [6] J. Gao, X. Wu, M. Xiao, V. P. Nguyen, L. Gao, T. Rabczuk, Multi-patch isogeometric topology optimization for cellular structures with flexible designs using nitsche's method, *Computer Methods in Applied Mechanics and Engineering* 410 (2023) 116036.
- [7] X. Zhang, M. Xiao, L. Gao, J. Gao, A t-splines-oriented isogeometric topology optimization for plate and shell structures with arbitrary geometries using bézier extraction, *Computer Methods in Applied Mechanics and Engineering* 425 (2024) 116929.
- [8] X. Zhang, L. Gao, M. Xiao, J. Gao, T-splines-based panel method for aerodynamic topology optimization of engineering shell structures using isogeometric analysis, *Computer Methods in Applied Mechanics and Engineering* 444 (2025) 118154.
- [9] D. Lin, L. Gao, J. Gao, The lagrangian-eulerian described particle flow topology optimization (pfto) approach with isogeometric material point method, *Computer Methods in Applied Mechanics and Engineering* 440 (2025) 117892.
- [10] J. Zhu, D. Lin, L. Gao, H. Cao, J. Gao, Particle-flow topology optimization with the sph kernel for hinge-free designs of compliant mechanisms using isogeometric material point method, *Chinese Journal of Mechanical Engineering* (2026) 100229.
- [11] O. Sigmund, J. Petersson, Numerical instabilities in topology optimization: a survey on procedures dealing with checkerboards, mesh-dependencies and local minima, *Structural optimization* 16 (1) (1998) 68–75.
- [12] G. I. Rozvany, A critical review of established methods of structural topology optimization, *Structural and multidisciplinary optimization* 37 (3) (2009) 217–237.
- [13] K.-T. Cheng, N. Olhoff, An investigation concerning optimal design of solid elastic plates, *International Journal of Solids and Structures* 17 (3) (1981) 305–323.
- [14] Y. Lam, S. Santhikumar, Automated rib location and optimization for plate structures, *Structural and multidisciplinary optimization* 25 (1) (2003) 35–45.
- [15] H. Lee, J. Yeo, K. Kong, D. Myeong, D. Jang, J. Lee, H. Choi, N. Kim, S. Ryu, Bayesian optimization of tailgate rib structures enhancing structural stiffness under manufacturing constraints of injection molding, *Journal of Manufacturing Processes* 134 (2025) 739–748.
- [16] X. Ding, K. Yamazaki, Stiffener layout design for plate structures by growing and branching tree model (application to vibration-proof design), *Structural and Multidisciplinary Optimization* 26 (1) (2004) 99–110.
- [17] Y. Cheng, F. He, X. Lv, W. Cai, On the role of generating textual description for design intent communication in feature-based 3d collaborative design, *Advanced Engineering Informatics* 39 (2019) 331–346.
- [18] K. Mustapha, A survey of emerging applications of large language models for problems in mechanics, product design, and manufacturing, *Advanced Engineering Informatics* 64 (2025) 103066.
- [19] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer, High-resolution image synthesis with latent diffusion models, in: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 10684–10695.
- [20] B. Poole, A. Jain, J. T. Barron, B. Mildenhall, Dreamfusion: Text-to-3d using 2d diffusion, *arXiv preprint arXiv:2209.14988* (2022).
- [21] F. Mazé, F. Ahmed, Diffusion models beat gans on topology optimization, in: *Proceedings of the AAAI conference on artificial intelligence*, Vol. 37, 2023, pp. 9108–9116.
- [22] G. Giannone, A. Srivastava, O. Winther, F. Ahmed, Aligning optimization trajectories with diffusion models for constrained design generation, *Advances in neural information processing systems* 36 (2023) 51830–51861.
- [23] J.-H. Bastek, W. Sun, D. M. Kochmann, Physics-informed diffusion models, *arXiv preprint arXiv:2403.14404* (2024).
- [24] Y. Kwon, N. Kang, Three-dimensional deep shape optimization with a limited dataset, *Engineering Applications of Artificial Intelligence* 159 (2025) 111504.
- [25] A. Bekbolat, S. Kurokawa, F. H. Kamaru Zaman, M. A. Samad Kamal, E. Shehab, M. H. Ali, Diffusion models for topology optimization in 3d printing applications, *Journal of Applied Physics* 137 (13) (2025).
- [26] S. Yoo, N. Kang, Deepwheel: Generating a 3d synthetic wheel dataset for design and performance evaluation, *Journal of Mechanical Design* 148 (5) (2026) 051702.
- [27] S. Hong, Y. Kwon, D. Shin, J. Park, N. Kang, Deepjeb: 3d deep learning-based synthetic jet engine bracket dataset, *Journal of Mechanical Design* 147 (4) (2025) 041703.
- [28] J. Kim, Y. Kwon, N. Kang, Deep generative design for mass production, *arXiv preprint arXiv:2403.12098* (2024).
- [29] Y. Wu, L. Wang, Y. Liu, Machine-learning driven differentiable concurrent multiscale topology optimization framework for truss structures, *Computer Methods in Applied Mechanics and Engineering* 460 (2026) 119135.
- [30] Z. Li, L. Wang, X. Zhang, Y. Liu, A level-set-oriented and problem-independent robust topology optimization strategy based on convolutional neural networks and uncertainty clustering, *Computer Methods in Applied Mechanics and Engineering* 455 (2026) 118894.

- [31] Z. Li, L. Wang, F. Li, Model-and-data-driven robust concurrent optimization for structural topology and device layout integrating load location and non-design domain effects, *Aerospace Science and Technology* (2025) 111566.
- [32] X. Zhao, L. Wang, Y. Liu, A novel reliability-oriented concurrent topology optimization framework based on improved sequential decoupling and interval-modeled boundary modification strategies, *Composite Structures* (2026) 120211.
- [33] W. Zhang, G. Zhao, L. Su, Research on multi-stage topology optimization method based on latent diffusion model, *Advanced Engineering Informatics* 63 (2025) 102966.
- [34] W. Zhang, K. Yu, L. Lu, L. Su, S. L. Sing, TO-MambaLDM: Multi-physics informed Mamba-enhanced latent diffusion for cooling structure topology optimization, *Advanced Engineering Informatics* 71 (2026) 104292.
- [35] T. Liu, F. Yu, Q. Li, X.-J. Wan, Z. Han, S. Liu, Topology optimization approach using a training-dataset-free neural network reparameterization framework, *Advanced Engineering Informatics* 70 (2026) 104111.
- [36] Y. Kwon, N. Kang, Three-dimensional geometric augmentation guided by engineering uncertainty, *SSRN preprint* (2025). doi:10.2139/ssrn.5380284.
- [37] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, R. Ng, Nerf: Representing scenes as neural radiance fields for view synthesis, *Communications of the ACM* 65 (1) (2021) 99–106.
- [38] B. T. Feng, J. Smith, M. Rubinstein, H. Chang, K. L. Bouman, W. T. Freeman, Score-based diffusion models as principled priors for inverse imaging, in: *Proceedings of the IEEE/CVF international conference on computer vision*, 2023, pp. 10520–10531.
- [39] H. Chung, D. Ryu, M. T. McCann, M. L. Klasky, J. C. Ye, Solving 3d inverse problems using pre-trained 2d diffusion models, in: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2023, pp. 22542–22551.
- [40] J. Song, A. Vahdat, M. Mardani, J. Kautz, Pseudoinverse-guided diffusion models for inverse problems, in: *International conference on learning representations*, 2023.
- [41] S. Zhong, P. Punpongsanon, D. Iwai, K. Sato, Topology optimization with text-guided stylization, *Structural and Multidisciplinary Optimization* 66 (12) (2023) 256.
- [42] Z. Wang, C. Fu, S.-f. Niu, S.-j. Hu, Research on bio-inspired product design based on knowledge graph and semantic fusion diffusion model, *Advanced Engineering Informatics* 62 (2024) 102797.
- [43] Z. Liang, Y.-F. Zhang, Y. Wang, W. Li, Integrating large models with topology optimization for conceptual design realization, *Advanced Engineering Informatics* 67 (2025) 103524.
- [44] B. Bourdin, Filters in topology optimization, *International journal for numerical methods in engineering* 50 (9) (2001) 2143–2158.
- [45] T. E. Bruns, D. A. Tortorelli, Topology optimization of non-linear elastic structures and compliant mechanisms, *Computer methods in applied mechanics and engineering* 190 (26–27) (2001) 3443–3459.
- [46] J. K. Guest, J. H. Prévost, T. Belytschko, Achieving minimum length scale in topology optimization using nodal design variables and projection functions, *International journal for numerical methods in engineering* 61 (2) (2004) 238–254.
- [47] R. Mindlin, Influence of rotatory inertia and shear on flexural motions of isotropic, elastic plates (1951).
- [48] E. Reissner, The effect of transverse shear deformation on the bending of elastic plates (1945).
- [49] K.-J. Bathe, E. N. Dvorkin, A four-node plate bending element based on mindlin/reissner plate theory and a mixed interpolation, *International Journal for Numerical Methods in Engineering* 21 (2) (1985) 367–383.
- [50] H. Rodrigues, P. Fernandes, A material based model for topology optimization of thermoelastic structures, *International Journal for Numerical Methods in Engineering* 38 (12) (1995) 1951–1965.
- [51] J. Ho, A. Jain, P. Abbeel, Denoising diffusion probabilistic models, *Advances in neural information processing systems* 33 (2020) 6840–6851.
- [52] J. Ho, T. Salimans, Classifier-free diffusion guidance, *arXiv preprint arXiv:2207.12598* (2022).
- [53] B. Tang, J. Wang, Z. Wu, L. Zhang, Stable score distillation for high-quality 3d generation, *arXiv preprint arXiv:2312.09305* (2023).
- [54] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980* (2014).
- [55] T. Y. Zhang, C. Y. Suen, A fast parallel algorithm for thinning digital patterns, *Communications of the ACM* 27 (3) (1984) 236–239.
- [56] D. McAllister, S. Ge, J.-B. Huang, D. W. Jacobs, A. A. Efros, A. Holynski, A. Kanazawa, Rethinking score distillation as a bridge between image distributions, *Advances in Neural Information Processing Systems* 37 (2024) 33779–33804.
- [57] K. Svanberg, The method of moving asymptotes—a new method for structural optimization, *International journal for numerical methods in engineering* 24 (2) (1987) 359–373.
- [58] C. Schuhmann, R. Beaumont, R. Vencu, C. Gordon, R. Wightman, M. Cherti, T. Coombes, A. Katta, C. Mullis, M. Wortsman, et al., Laion-5b: An open large-scale dataset for training next generation image-text models, *Advances in neural information processing systems* 35 (2022) 25278–25294.
- [59] F. Wang, B. S. Lazarov, O. Sigmund, On projection methods, convergence and robust formulations in topology optimization, *Structural and multidisciplinary optimization* 43 (6) (2011) 767–784.
- [60] H. Jun, A. Nichol, Shap-e: Generating conditional 3d implicit functions, *arXiv preprint arXiv:2305.02463* (2023).
- [61] Y. Shi, P. Wang, J. Ye, M. Long, K. Li, X. Yang, Mvdream: Multi-view diffusion for 3d generation, *arXiv preprint arXiv:2308.16512* (2023).
- [62] J. K. Guest, Imposing maximum length scale in topology optimization, *Structural and Multidisciplinary Optimization* 37 (2009) 463–473.
- [63] S. Yoo, S. Lee, S. Kim, K. H. Hwang, J. H. Park, N. Kang, Integrating deep learning into cad/cae system: generative design and evaluation of 3d conceptual wheel: Integrating deep learning into cad/cae system: Generative design and evaluation

of 3d conceptual wheel, Structural and multidisciplinary optimization 64 (4) (2021) 2725–2747.