

CoPhi: Conditionnement Physique de la génération 3D

I Présentation du projet

Les modèles génératifs produisent aujourd’hui des images, des vidéos et des objets 3D d’un réalisme saisissant. Cependant, ils privilégient la qualité visuelle sans garantir la cohérence physique des contenus générés. Cette limite freine leur utilisation pour les applications industrielles, comme la conception, la robotique ou les systèmes autonomes, pour lesquelles les objets produits doivent respecter des propriétés vérifiables.

Prévu pour une période de trois ans au CEA List, au sein du département DIASI et du service SIALV (laboratoire LVA), CoPhi s’inscrit dans un environnement scientifique comprenant notamment le projet Carnot Gen ϕ sur la génération 3D consciente de la physique. À partir d’une consigne telle que « génère une chaise en bois pouvant supporter une charge de 110 kg », ou d’un cahier des charges de contraintes physiques, le modèle devra produire un objet aux propriétés physiques cohérentes dont la validité pourra être vérifiée. Dans ce document, nous dénotons ces informations sous le terme de « **caractéristiques et contraintes physiques (CCP)** ».

Nous faisons l’hypothèse qu’une *représentation commune et structurée des formes et des CCP permettra de dépasser les approches actuelles et d’améliorer le réalisme physique des méthodes génératives 3D*. La validation de cette hypothèse nécessite de lever **trois verrous majeurs** :

- la rareté des données associant objets 3D et informations physiques ;
- l’absence de représentation permettant de composer et de manipuler des CCP hétérogènes ;
- l’intégration des contraintes dans la génération sans dégrader qualité et cohérence sémantiques.

Le projet poursuit en conséquence **trois objectifs scientifiques** :

- construire les données et le protocole nécessaires à l’apprentissage et à l’évaluation des CCP ;
- apprendre une représentation commune de l’objet physique et de son cahier des charges ;
- faire de cette représentation une nouvelle modalité de contrôle pour la génération 3D.

CoPhi bénéficiera, pour sa réalisation, de synergies naturelles avec le projet Carnot Gen ϕ . Gen ϕ établit un cadre général pour une génération 3D consciente de la physique, intégrant conjointement des contraintes géométriques, structurelles et physiques. Il prévoit notamment les premières briques de conditionnement du générateur. Ses travaux sur le contrôle structurel et la préparation des objets à la simulation fourniront des briques scientifiques et logicielles directement mobilisables par CoPhi, notamment pour produire les corpus et le benchmark de son premier lot. Cette complémentarité permettra à CoPhi de concentrer son ambition sur une représentation commune, composable et manipulable des propriétés matérielles, des sollicitations et des réponses physiques. Ces caractéristiques pourront être globales ou locales et décrire les comportements propres aux différentes parties d’un objet composé.

Face au coût des modèles génératifs, CoPhi privilégiera l’adaptation de modèles 3D préentraînés et la réutilisation de leurs représentations. Cette démarche, en cohérence avec les priorités du Programme et Équipements Prioritaires de Recherche en intelligence artificielle (PEPR IA), contribuera à rapprocher la génération 3D des modèles du monde physique, en donnant accès non seulement à l’apparence des objets, mais aussi à leurs matériaux et à leur comportement.

II État de l’art

Limites physiques de la génération 3D. Les modèles génératifs transforment désormais une consigne textuelle ou visuelle en images [24], vidéos [29], objets 3D [26, 33] ou environnements interactifs [19]. TRELLIS [33] et TRELLIS.2 [32] produisent des géométries et des apparences matérielles réalistes en transférant des modèles de fondation visuels vers la 3D [18, 25, 31]. Ce transfert pallie la rareté des données 3D, mais il reste à déterminer quelles propriétés physiques ces représentations encodent et comment les enrichir pour satisfaire un cahier des charges [1]. Les premières approches dépassant le réalisme visuel introduisent des structures de contrôle, telles que des hiérarchies ou des graphes [11, 17, 8, 35], ou des articulations [27, 13, 7]. CoPhi ciblera les propriétés matérielles et les réponses aux sollicitations en s’appuyant sur la simulation, Gen ϕ traitant les questions de structure et de simulabilité.

Génération guidée par la simulation physique. La simulation a été utilisée pour améliorer la stabilité mécanique d’objets générés [16, 15, 3, 22] ou optimiser leur structure interne [34, 14], y compris pendant la génération [12]. D’autres travaux se concentrent sur des applications spécialisées : TripOptimizer [28] et PhysGen [36] optimisent l’aérodynamique de géométries automobiles [5], tandis que DiffuseBot [30] génère des préhenseurs souples à partir de contraintes fonctionnelles. Si l’approche est prometteuse, ces méthodes restent toutefois limitées à une propriété physique et à une catégorie d’objets prédéfinis, avec une qualité et une diversité inférieures à celles des grands modèles génératifs 3D.

III Ambition du projet

L’ambition de CoPhi est de proposer une génération 3D, physique, réaliste et généraliste. Plutôt que de se focaliser sur une propriété, un simulateur ou une catégorie d’objets prédéfinis, le projet cherchera à représenter plusieurs caractéristiques et contraintes physiques dans un espace commun aux formes 3D et au langage, et à conditionner la génération 3D par cette représentation.

Avancées scientifiques. Le projet permettra d’apprendre une représentation commune reliant forme, matériau et réponse aux sollicitations physiques, exprimable aussi bien par une géométrie annotée que par le langage naturel. Il étudiera comment composer ces informations, manipuler la forme dans cet espace et contrôler directement un générateur 3D généraliste.

Sauts technologiques. CoPhi passera de modèles optimisant une propriété dans un scénario prédéfini à un démonstrateur capable de générer, modifier et vérifier un objet selon plusieurs exigences physiques. Il apportera ainsi une brique essentielle aux futurs assistants de conception et, plus largement, aux modèles capables de représenter le monde physique.

IV Approche proposée

Dans la suite de ce projet, nous désignons par **caractéristiques et contraintes physiques (CCP)** l’ensemble des informations relatives à la **géométrie**, à la **structure**, aux **matériaux** et au **comportement physique** d’un objet. Une CCP peut être globale ou locale et prendre la forme d’une propriété, d’une valeur cible ou de la réponse de l’objet à une sollicitation. Le tableau 1 en propose une taxonomie simplifiée.

Géométrie et topologie	Propriété du matériau	Réponse à une sollicitation	Stabilité et faisabilité
Symétrie, dimensions, épaisseur, volume, connexité, composantes	Densité, module d'Young, coefficient de Poisson, viscosité, frottement	Énergie de déformation, coefficients de traînée ou de portance	Stabilité, caractère autoportant, constructibilité

TABLEAU 1 – Taxonomie (simplifiée) des CCP envisagées, avec quelques exemples.

La réalisation du projet nécessite trois lots de travaux complémentaires. Le lot 1, « *Données physiques et benchmark multi-fidélité* », fournira les données et le protocole d'évaluation. Le lot 2, « *Représentation commune des formes et des CCP* », apprendra à représenter et à manipuler conjointement l'objet et son cahier des charges. Enfin, le lot 3, « *Génération 3D conditionnée par les CCP* », intégrera cette représentation comme modalité de contrôle d'un modèle génératif 3D.

Ressources. La réalisation du projet reposera sur le porteur, le recrutement d'un doctorant pour trois ans et l'accueil d'un stagiaire, principalement mobilisé sur la constitution des données du lot 1. Deux stations de travail équipées pour l'apprentissage 3D seront nécessaires, l'une pour le porteur et l'autre pour le doctorant. Des frais seront également prévus pour quatre missions scientifiques consacrées à la présentation des résultats et au développement des collaborations.

IV.1 Lot 1 — Données physiques et benchmark multi-fidélité

Verrou scientifique et objectif. Les données associant des formes 3D et des propriétés physiques *restent rares et spécialisées*. Ce lot vise à déverrouiller la situation pour constituer le socle expérimental nécessaire à CoPhi. Il ciblera en priorité les réponses à des sollicitations — dans la continuité des travaux existants [36, 20] — ainsi que les propriétés mécaniques des matériaux.

Tâche 1.1. Corpus de données physiques de grande échelle. Dans cette première tâche, nous construirons une base de données associant des géométries 3D diversifiées à leurs CCP. Nous développerons un pipeline reproductible permettant de paramétrer et de simuler ces géométries à l'aide de moteurs physiques existants [6]. Les modèles 3D proviendront de créations artistiques [4] ou de générateurs procéduraux [23], et seront complétés par des descriptions textuelles [33]. Ce corpus sera aussi enrichi par des données existantes [36].

Tâche 1.2. Données de référence expert. Nous constituerons un corpus de plus petite échelle réunissant des géométries sélectionnées, des descriptions textuelles détaillées et des CCP obtenues dans des conditions précisément documentées. Construit avec les experts en simulation du CEA List, notamment du LSI et son expertise d'apprentissage pour la simulation mécanique [21], ce jeu de référence permettra d'évaluer la précision physique et la généralisation des modèles.

Résultats et livrables. Ce lot aboutira à *un corpus de grande échelle formes-CCP* destiné à l'apprentissage – et son pipeline reproductible de génération et de simulation – et à *un benchmark comprenant un jeu de référence expert et son protocole d'évaluation*. Les deux livrables feront l'objet d'une publication scientifique ; les données et logiciels diffusables seront rendus publics.

Calendrier, dépendances et ressources. Engagé dès M1, ce lot livrera en un an le socle expérimental de CoPhi. Indépendant des autres lots, il produira néanmoins les données et le protocole d'évaluation exploités dans les lots suivants. Le porteur mènera ces travaux avec un doctorant et un stagiaire consacré à la production des données. Le socle expérimental de Gen ϕ servira de base de travail pour la base d'apprentissage et une collaboration avec les experts en simulation du CEA List est envisagée pour la base experte.

IV.2 Lot 2 – Représentation commune des formes et des CCP

Verrou scientifique et objectif. Les CCP regroupent des informations hétérogènes – *géométrie, matériaux, réponses physiques* – que les représentations 3D actuelles *ne permettent pas* d’exprimer ou de manipuler. Ce lot fera des CCP *une nouvelle modalité de contrôle*, apprise conjointement avec les formes. L’objectif est de donner une représentation commune à l’objet physique et à son cahier des charges, fournis sous la forme d’une géométrie annotée ou décrits en langage naturel.

Tâche 2.1. Représentation conjointe des formes et des CCP. Notre candidat privilégié est de partir de l’espace latent de TRELIS.2 [32], dont les représentations spatialement alignées de la géométrie et de l’apparence matérielle seront enrichies avec les CCP du lot 1. Nous explorerons également le transfert de modèles de fondation 2D vers la physique, dans le style du transfert de modèles 2D vers la 3D [33, 9], ainsi que l’inversion de modèles génératifs 3D [2].

Tâche 2.2. Manipulation et optimisation dans l’espace latent. Nous développerons un démonstrateur permettant d’encoder une forme existante, de modifier une ou plusieurs CCP dans sa représentation, puis de décoder la forme optimisée. Ce démonstrateur validera le caractère manipulable de la représentation, dans la continuité de travaux récents sur l’optimisation de formes à partir de représentations physiques apprises [36, 28].

Tâche 2.3. Alignement avec les cahiers des charges en langage naturel. Nous apprendrons à convertir un cahier des charges textuel en une représentation de ses CCP, alignée avec celle d’une forme 3D annotée. Nous nous appuierons notamment sur les descriptions expertes du lot 1 et sur notre expérience de l’alignement entre texte et représentations 3D [9].

Résultats et livrables. Le livrable comprendra *une représentation pré-entraînée* permettant d’estimer, de composer et de manipuler les CCP d’une forme annotée ou d’un cahier des charges textuel. Elle sera accompagnée d’un démonstrateur d’optimisation et évaluée sur le benchmark du lot 1. Les tâches 2.1 et 2.2 seront valorisées par une publication scientifique et la tâche 2.3 fera l’objet d’une publication dédiée. Le code et les modèles diffusables seront rendus publics.

Calendrier, dépendances et ressources. Les premières études sur les représentations préentraînées pourront commencer dès M6, mais l’apprentissage et l’évaluation du modèle dépendront de ses données et protocoles. Le cœur du lot sera mené de M12 à M24 en travail conjoint entre le doctorant et le porteur du projet, et en collaboration avec les équipes du CEA List concernées.

IV.3 Lot 3 – Génération 3D conditionnée par les CCP

Verrou scientifique et objectif. La génération 3D actuelle peut suivre une intention visuelle ou textuelle, *mais pas un cahier des charges physique hétérogène*. Nous conditionnerons un modèle génératif par un cahier des charges en utilisant notre représentation commune, tout en préservant la qualité, la diversité et la cohérence des formes produites.

Tâche 3.1. Adaptation d’un générateur 3D pré-entraîné. Nous adapterons un modèle pré-entraîné [33, 32] à la condition développée dans le lot 2. Le corpus à grande échelle du lot 1 permettra d’apprendre ce nouveau conditionnement, tandis que des mécanismes d’adaptation paramétriquement efficaces [10] préserveront les capacités géométriques du modèle initial.

Tâche 3.2. Démonstration et vérification physique. Nous construirons un démonstrateur de génération à partir de cahiers des charges. Le benchmark du lot 1 permettra de vérifier la satisfaction des contraintes et la qualité des objets produits. Nous testerons la généralisation à des combinaisons inédites et la capacité du modèle à s’adapter aux scénarios physiquement invalides.

Résultats et livrables. Ce lot livrera la capacité finale du projet : *générer un objet 3D à partir de son intention - visuelle ou textuelle - et de son cahier des charges physique*. Un modèle, un démonstrateur et une évaluation quantitative matérialiseront cette avancée. Une publication scientifique, le code et les poids diffusables accompagneront ces résultats.

Calendrier, dépendances et ressources. Les premières preuves de concept pour l’adaptation de TRELLIS pourront commencer à M18, en coordination avec les résultats du projet Carnot Gen ϕ . De M24 à M36, ce lot réunira les données du lot 1 et la représentation du lot 2 pour produire le démonstrateur final de CoPhi. Le porteur du projet et le doctorant conduiront cette intégration, avec l’appui des membres du projet Gen ϕ .

V Liens avec le PEPR IA

CoPhi répond tout d’abord à un enjeu de frugalité en données et en calcul. Malgré l’existence de grandes collections d’objets 3D [4], peu de propriétés physiques sont documentées. Le projet privilégiera donc l’adaptation de modèles préentraînés afin de réduire conjointement les besoins en données et en calcul de l’apprentissage. Ces travaux renforceront les collaborations du CEA List, et notamment du SIALV, au sein des projets **SHARP** – sur l’apprentissage à partir de données limitées et hétérogènes – et **HOLIGRAIL** – sur la réduction des coûts d’entraînement et d’inférence.

CoPhi répond également à des problématiques d’IA de confiance. En représentant explicitement les propriétés physiques et en vérifiant leur satisfaction, le projet rendra les modèles génératifs plus contrôlables et leurs performances quantifiables. Le projet ouvrira ainsi une nouvelle voie d’implication du CEA dans le PEPR IA autour de **Géné-Pi**, dont les recherches sur la fiabilité des modèles génératifs et l’intégration de contraintes physiques sont complémentaires du cadre applicatif développé par CoPhi.

VI Liens avec le Département DIASI

Ce projet de recherche s’inscrit pleinement dans les orientations stratégiques du CEA List et de son département DIASI en matière d’intelligence artificielle générative. Principalement rattaché au laboratoire LVA, il associera également les laboratoires LASTI et LSI.

Le LASTI, le LVA et le LSI sont déjà partenaires au sein du projet Carnot Gen ϕ , et comme expliqué, de fortes émulations se mettront en place entre ces deux projets. La thèse de E. Louvard notamment, a débuté sur le conditionnement structurel de la génération 3D.

Au-delà du projet Gen ϕ , le LASTI et le LVA collaborent régulièrement sur les modèles génératifs de textes, d’images, de vidéos et d’objets 3D. Le LVA a ainsi travaillé sur des approches de génération d’images et de vidéos conditionnées spatialement (stage et thèse de A. Sangare) et sur la génération d’objets 3D sous forme de Gaussian Splatting (thèse de B. Engel). De son côté, le LASTI travaille aussi sur les modèles génératifs visuels, en particulier sur leur contrôle (thèses de A. Plumerault, P. Doubinsky et P. Grimal) et leur explicabilité (thèse de C. Cornet).

Des liens seront aussi tissés avec le LSI qui développe la plateforme logicielle 2MORDOR (MODèle de MONde pour la Recherche et le Déploiement Opérationnel de Robots intelligents). Cette plateforme intègre une représentation du monde (géométrique, sémantique et physique) pour des applications de robotique industrielle intelligente.

Références

- [1] Florian Bordes, Quentin Garrido, Justine T Kao, Adina Williams, Michael Rabbat, and Emmanuel Dupoux. Intphys 2 : Benchmarking intuitive physics understanding in complex synthetic environments, 2025.
- [2] Victoria Yue Chen, Emery Pierson, Léopold Maillard, and Maks Ovsjanikov. Beyond prompts : Unconditional 3D inversion for out-of-distribution shapes. In *Proceedings of the European Conference on Computer Vision*, 2026.
- [3] Yunuo Chen, Tianyi Xie, Zeshun Zong, Xuan Li, Feng Gao, Yin Yang, Ying Nian Wu, and Chenfanfu Jiang. Atlas3D : Physically constrained self-supporting text-to-3D for simulation and fabrication. In *Advances in Neural Information Processing Systems*, volume 37, 2024.
- [4] Matt Deitke, Ruoshi Liu, Matthew Wallingford, Huong Ngo, Oscar Michel, Aditya Kusupati, Alan Fan, Christian Laforte, Vikram Voleti, Samir Yitzhak Gadre, Eli VanderBilt, Aniruddha Kembhavi, Carl Vondrick, Georgia Gkioxari, Kiana Ehsani, Ludwig Schmidt, and Ali Farhadi. Objaverse-XL : A universe of 10m+ 3D objects. In *Advances in Neural Information Processing Systems*, volume 36, 2023.
- [5] Mohamed Elrefaie, Florin Morar, Angela Dai, and Faez Ahmed. Drivaernet++ : A large-scale multimodal car dataset with computational fluid dynamics simulations and deep learning benchmarks. *Advances in Neural Information Processing Systems*, 37 :499–536, 2024.
- [6] François Faure, Christian Duriez, Hervé Delingette, Jérémie Allard, Benjamin Gilles, Stéphanie Marchesseau, Hugo Talbot, Hadrien Courtecuisse, Guillaume Bousquet, Igor Peterlik, and Stéphane Cotin. SOFA : A Multi-Model Framework for Interactive Physical Simulation. In Yohan Payan, editor, *Soft Tissue Biomechanical Modeling for Computer Assisted Surgery*, volume 11 of *Studies in Mechanobiology, Tissue Engineering and Biomaterials*, pages 283–321. Springer, June 2012.
- [7] Daoyi Gao, Yawar Siddiqui, Lei Li, and Angela Dai. MeshArt : Generating articulated meshes with structure-guided transformers. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2025.
- [8] Lin Gao, Jie Yang, Tong Wu, Yu-Jie Yuan, Hongbo Fu, Yu-Kun Lai, and Hao Zhang. SDM-NET : Deep generative network for structured deformable mesh. *ACM Transactions on Graphics*, 38(6), 2019.
- [9] Souhail Hadgi, Bingchen Gong, Ramana Sundararaman, Emery Pierson, Lei Li, Peter Wonka, and Maks Ovsjanikov. Patchalign3d : Local feature alignment for dense 3d shape understanding, 2026.
- [10] Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA : Low-rank adaptation of large language models. In *Proceedings of the International Conference on Learning Representations*, 2022.
- [11] Jun Li, Kai Xu, Siddhartha Chaudhuri, Ersin Yumer, Hao Zhang, and Leonidas Guibas. GRASS : Generative recursive autoencoders for shape structures. *ACM Transactions on Graphics*, 36(4), 2017.
- [12] Daochang Liu, Junyu Zhang, Anh-Dung Dinh, Eunbyung Park, Shichao Zhang, Ajmal Mian, Mubarak Shah, and Chang Xu. Generative physical AI in vision : A survey, 2025.

- [13] Jiayi Liu, Hou In Ivan Tam, Ali Mahdavi-Amiri, and Manolis Savva. CAGE : Controllable articulation generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024.
- [14] Rundong Luo, Haoran Geng, Congyue Deng, Puhao Li, Zan Wang, Baoxiong Jia, Leonidas Guibas, and Siyuan Huang. PhysPart : Physically plausible part completion for interactable objects. In *Proceedings of the IEEE International Conference on Robotics and Automation*, 2025.
- [15] Mariem Mezghanni, Théo Bodrito, Malika Boulkenafed, and Maks Ovsjanikov. Physical simulation layer for accurate 3D modeling. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.
- [16] Mariem Mezghanni, Mohamed Boulkenafed, André Lieutier, and Maks Ovsjanikov. Physically-aware generative network for 3D shape modeling. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021.
- [17] Kaichun Mo, Paul Guerrero, Li Yi, Hao Su, Peter Wonka, Niloy J. Mitra, and Leonidas J. Guibas. StructureNet : Hierarchical graph networks for 3D shape generation. *ACM Transactions on Graphics*, 38(6), 2019.
- [18] Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, et al. Dinov2 : Learning robust visual features without supervision. *arXiv preprint arXiv :2304.07193*, 2023.
- [19] Jack Parker-Holder and Shlomi Fruchter. Genie 3 : A new frontier for world models. Google DeepMind, August 2025. Online article.
- [20] Emery Pierson, Mohamed Daoudi, and Alice-Barbara Tumpach. A riemannian framework for analysis of human body surface. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 2991–3000, 2022.
- [21] Louen Pottier, Anders Thorin, and Francisco Chinesta. Latent-energy-based NNs : An interpretable neural network architecture for model-order reduction of nonlinear statics in solid mechanics. *Journal of the Mechanics and Physics of Solids*, 194 :105953, 2024.
- [22] Ava Pun, Kangle Deng, Ruixuan Liu, Deva Ramanan, Changliu Liu, and Jun-Yan Zhu. Generating physically stable and buildable brick structures from text. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2025.
- [23] Alexander Raistrick, Lingjie Mei, Karhan Kayan, David Yan, Yiming Zuo, Beining Han, Hongyu Wen, Meenal Parakh, Stamatis Alexandropoulos, Lahav Lipson, Zeyu Ma, and Jia Deng. Infinigen indoors : Photorealistic indoor scenes using procedural generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 21783–21794, June 2024.
- [24] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.
- [25] Oriane Siméoni, Huy V Vo, Maximilian Seitzer, Federico Baldassarre, Maxime Oquab, Cijo Jose, Vasil Khalidov, Marc Szafraniec, Seungeun Yi, Michaël Ramamonjisoa, et al. Dinov3. *arXiv preprint arXiv :2508.10104*, 2025.
- [26] Tencent Hunyuan3D Team. Hunyuan3D 2.0 : Scaling diffusion models for high-resolution textured 3D assets generation, 2025.

- [27] Xiaoyu Tian, Yijia Yang, and Qi Wu. ShapeScaffolder : Structure-aware 3D shape generation from text. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023.
- [28] Parsa Vatani, Mohamed Elrefaie, Farhad Nazarpour, and Faez Ahmed. TripOptimizer : Generative three-dimensional shape optimization and drag prediction using triplane variational autoencoder networks. *Physics of Fluids*, 37(12), 2025.
- [29] Wan Team. Wan : Open and advanced large-scale video generative models, 2025.
- [30] Tsun-Hsuan Johnson Wang, Pingchuan Liu, Xiaohang Zhang, and Wojciech Matusik. Diffuse-Bot : Breeding soft robots with physics-augmented generative diffusion models. In *Advances in Neural Information Processing Systems*, volume 36, 2023.
- [31] Thomas Wimmer, Peter Wonka, and Maks Ovsjanikov. Back to 3d : Few-shot 3d keypoint detection with back-projected 2d features. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4154–4164, 2024.
- [32] Jianfeng Xiang, Xiaoxue Chen, Sicheng Xu, Ruicheng Wang, Zelong Lv, Yu Deng, Hongyuan Zhu, Yue Dong, Hao Zhao, Nicholas Jing Yuan, and Jiaolong Yang. Native and compact structured latents for 3D generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2026.
- [33] Jianfeng Xiang, Zelong Lv, Sicheng Xu, Yu Deng, Ruicheng Wang, Bowen Zhang, Dong Chen, Xin Tong, and Jiaolong Yang. Structured 3d latents for scalable and versatile 3d generation. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 21469–21480, 2025.
- [34] Qingshan Xu, Jiao Liu, Melvin Wong, Caishun Chen, and Yew-Soon Ong. Precise-physics driven text-to-3D generation, 2024.
- [35] Jie Yang, Kaichun Mo, Yu-Kun Lai, Leonidas J. Guibas, and Lin Gao. DSG-Net : Learning disentangled structure and geometry for 3D shape generation. *ACM Transactions on Graphics*, 42(1), 2022.
- [36] Yingxuan You, Chen Zhao, Hantao Zhang, Ming Xu, and Pascal Fua. PhysGen : Physically grounded 3D shape generation for industrial design. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2026.