Physical Design using Differentiable Learned Simulators

Anonymous Author(s)
Affiliation
Address
email

Abstract

Designing physical artifacts that serve a purpose—such as tools and other functional structures—is central to engineering as well as everyday human behavior. Though automating design using machine learning has tremendous promise, existing methods are often limited by the task-dependent distributions they were exposed to during training. Here we showcase a *task-agnostic* approach to inverse design, by combining general-purpose learned simulators with gradient-based design optimization. This constitutes a simple, fast, and reusable approach that solves high-dimensional problems with complex physical dynamics, including designing surfaces and tools to manipulate fluid flows and optimizing the shape of an airfoil to minimize drag. This framework produces high-quality designs by propagating gradients through trajectories of hundreds of steps, even when using models that were pre-trained for single-step predictions on data substantially different from the design tasks. In our fluid manipulation tasks, the resulting designs outperformed those found by sampling-based optimization techniques. In airfoil design, they matched the quality of those obtained with a specialized solver. Our results suggest that despite some remaining challenges, machine learning-based simulators are maturing to the point where they can support general-purpose design optimization across a variety of domains.

1 Introduction

Automatically designing objects to exhibit a desired property—often referred to as *inverse design*—has the potential to transform science and engineering, including aerodynamics [27], material design [14], optics [21], and robotics [34, 72]. Despite its promise, widespread practice of inverse design has been limited by the availability of fast, general-purpose forward models. Traditionally, many methods rely on specialized "classical" solvers, which are handcrafted to simulate a particular physical process. However, these models can sometimes be slow, may not provide gradients, or in certain domains, might even be unknown [23].

To mitigate these issues, machine learning is playing an increasingly important role by substituting all or part of the inverse design pipeline. "End-to-end" methods can bypass expensive simulation and data collection by learning a mapping directly from design parameters (such as the shape of an airplane wing) to a task objective (such as the observed drag) [57, 73, 19]. Learned generative models or policies can propose candidate design parameters for a given task [50, 17, 19, 24, 12, 45]. These types of methods can be used to solve narrowly defined tasks more efficiently than traditional methods, but come with significant limitations: they require access to a dataset of example designs and tasks for training, and often cannot generalize to tasks or designs outside this distribution [69, 74].

In this work, we focus on a *task-agnostic* approach to inverse design, by employing learned simulator models pre-trained on dynamics with *no* access to the task objective or design space during training,

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

and that, once trained, can be used to solve many new design tasks. This approach, however, requires models with sufficient long-term stability and strong generalization, which commonly applied black-box model architectures such as CNNs often lack [43, 67, 70, 4].

A class of learned physics simulators based on graph neural networks (GNNs) [55, 61] could be a promising candidate for supporting inverse design, as they have shown success in general-purpose physical prediction, exhibiting high accuracy and generalization for a broad range of physics tasks. However, these models have so far mostly been tested for forward prediction alone. To be useful in a design setting, they must also be smooth and robust, as optimizers will exploit unreliable regions in the state space [51]. Moreover, to use scalable gradient-based optimization techniques, simulation models must additionally exhibit gradient stability over extremely large state and action spaces, and over hundreds of time-steps. It is unclear if graph network simulators exhibit these properties.

In this paper, we investigate whether learned GNN simulators can support effective inverse design. We pretrain a GNN simulator on a task-independent dynamics dataset, and at test time perform gradient-based optimization for zero-shot design across a variety of tasks. We perform a large set of experiments and find that (1) the learned simulator model is accurate and flexible enough to support a wide variety of design tasks; (2) it generalizes well enough to support design tasks far outside the training regime; (3) gradients from the learned model turn out to be very useful for inverse design, with gradient-based optimization strongly outperforming a variety of sampling-based optimizers; (4) and surprisingly, gradients remain stable even when rolling the model out for hundreds of steps. In experiments on high-dimensional fluid manipulation (2D FLUID TOOLS and 3D WATERCOURSE) we demonstrate that gradient-based optimization on learned models can find high-quality designs over hundreds of time steps, through states with thousands of particles, in tasks with up to 625 design parameters. The same model could match performance of a specialized solver on aerodynamics design (AIRFOIL) with a much simpler setup, highlighting the flexibility and precision of this approach.

2 Related Work

Modern design techniques rely on physical simulators so that designs can be evaluated virtually before being tested physically. **Solving inverse design problems with simulators** has been studied in a variety of disciplines, including nanophotonics [52], material science [25], mechanical design [22], and aerodynamics [2, 58]. Many of these rely on black-box optimization, including active and adaptive sampling methods [66, 3]. In recent years, differentiable simulators [29, 42, 62] have garnered attention, as they allow for more sample-efficient gradient-based optimization. However, like classical solvers, they are typically narrow in application scope, as many simulation techniques are hard to express as a differentiable program (e.g. constraint dynamics and multiphysics coupling). Our work aims to lessen this burden by learning these simulators from data.

Machine learning for inverse design replaces or augments the simulation model and/or the optimization procedure with learned modules [e.g. 16, 19, 33, 28, 36, 40, 46, 48, 50, 57, 64, 75]. Generative approaches enable efficient optimization by learning to sample more likely designs (thus calling an expensive simulator fewer times). End-to-end predictive approaches learn a surrogate model that maps designs to task rewards directly (replacing the simulator entirely). This permits not only sampling-based optimization, but also gradient-based techniques when the surrogate model is differentiable [57]. However, surrogate models often generalize poorly to designs that lie outside of their training dataset, so many ML approaches for design must keep the optimization procedure close to the training dataset [12, 30], or introduce alternative objectives for the surrogate model to increase robustness [69, 74]. In all cases, these methods require access to datasets with (design, task reward) pairs on the specific task, making zero-shot generalization to new tasks impossible. Instead, we propose to learn the underlying physical dynamics *independent of the design task*, allowing both the design space and task to be specified at design time.

In this way, our approach is related to methods in **model-based reinforcement learning** (MBRL), which learns action-conditioned transition models to support flexible downstream planning and control. However, learned simulators are rarely accurate, stable, and generalizable enough to support better downstream task performance [35, 51]. Techniques from MBRL need to mitigate model error either by learning a global, task-conditioned *policy* during model learning [43] (which then cannot be adapted to novel tasks on-the-fly), or by *re-planning* at every timestep to avoid error accumulation [15]. Because of such sensitivities, these methods have only been demonstrated in

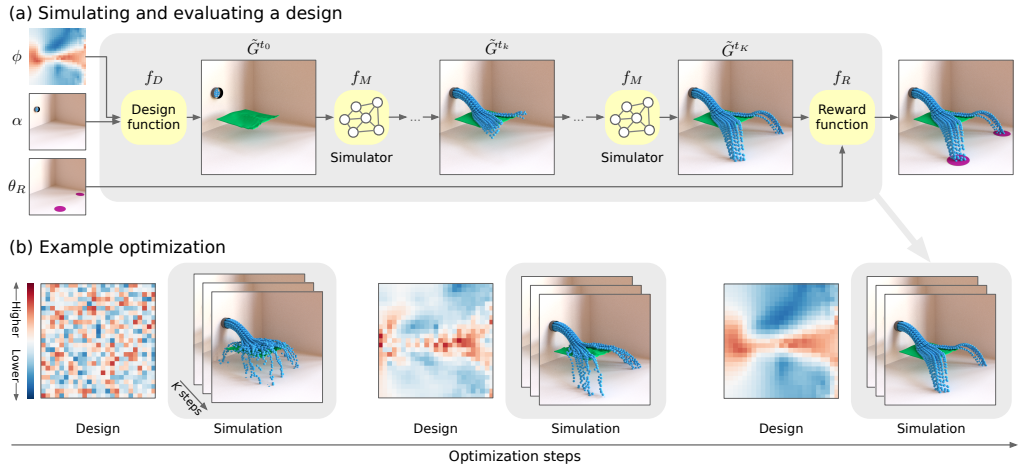

Figure 1: Optimizing a physical design. Here, the goal is to direct a stream of water (shown in blue) into two "pools" (shown in purple) by designing a "landscape" (shown in green) parameterized as a 2D height field. (a) The simulation pipeline takes in a design ϕ and initial conditions α and uses the design function f_D to produce an initial state. The simulation is rolled out with a pre-trained learned simulator f_M for K steps, at which point the final state is passed (along with reward parameters θ_R) to the reward function f_R, which computes the quality of the design. (b) Each step of optimization involves rolling out the simulation and then adjusting the design (ϕ) accordingly using an optimizer such as gradient descent or CEM. Shown are selected frames from gradient-based optimization in the *2 Pools* task of the 3D WATERCOURSE domain.

relatively small-scale problems, for example with (up to 21 degrees of freedom DOF [10, 60, 71]) or more complex domains with small action spaces (6 DOF [49]), unlike the thousands of state dimensions and hundreds of action dimensions and timesteps we consider here.

We focus on **graph neural networks** as our simulator model, as they have recently demonstrated high accuracy, stability, domain-generality, and generalization performance in the forwards direction for physics domains [7, 55, 61]. However, it is not known whether these simulators can now support gradient-based, high-dimensional inverse design which demands high accuracy, well-behaved gradients, long-term rollout stability, and generalization beyond the training data. Our experimental investigations show that they can, and that particular architectural choices are critical to this success.

3 Problem Formulation

Consider the design task depicted in Figure 1, in which the goal is to direct a stream of water (shown in blue) into two "pools" (shown in purple) by designing a "landscape" (shown in green) parameterized as a 2D height field. Here, an ideal design will create ridges and valleys that direct fluid into the two targets. In the next sections, we formalize what it means to find and evaluate such a design and discuss our choices for simulator and optimizer.

Learned simulators To demonstrate the utility of learned simulators for finding physical designs, we rely on the recently developed MESHGRAPHNETS model [55], which is an extension of the GNS model for particle simulation [61]. MESHGRAPHNETS is a type of message-passing graph neural network (GNN) that performs both edge and node updates [6, 32], and which was designed specifically for physics simulation by enforcing locality and translation equivariance. We briefly summarize the learned simulator here, and refer interested readers to the original papers for details.

We consider simulations over physical states represented as graphs $G \in \mathcal{G}$. The state $G = (V, E)$ has nodes V connected by edges E, where each node $v \in V$ is associated with a position $\mathbf{u}_v$ and additional dynamical quantities $\mathbf{q}_v$. These graphs may be either meshes (as in MESHGRAPHNETS) or particle systems (as in GNS). In a mesh-based system (such as AIRFOIL), V and E correspond to vertices and edges in the mesh, respectively. In a particle system (such as 2D FLUID TOOLS), each node corresponds to a particle and edges are computed dynamically based on proximity. Under this framework, we can also consider hybrid mesh-particle systems (such as 3D WATERCOURSE).

The model is constrained to perform local computation, by performing message passing over these edges, and to perform equivariant computations by encoding position as the relative displacement and distance between neighboring particles. Intuitively, this allows it to generalize well to scenarios where the local interactions are the same but the global arrangement of particles changes. See Appendix B for model implementation details, and Appendix C for domain-specific details.

The simulation dynamics are given by a "ground-truth" simulator $f_S : \mathcal{G} \to \mathcal{G}$ which maps the state at time t to that at time $t + \Delta t$. The simulator f_S can be applied iteratively over K time steps to yield a trajectory of states, or a "rollout," which we denote $(G^{t_0}, ..., G^{t_K})$. Using MESHGRAPHNETS, we learn an approximation f_M of the ground-truth simulator f_S. The learned simulator f_M can be similarly applied to produce rollouts $(\tilde{G}^{t_0}, \tilde{G}^{t_1}, ..., \tilde{G}^{t_K})$, where $\tilde{G}^{t_0} = G^{t_0}$ represents initial conditions given as input. We note that a learned simulator allows us to take much larger time-steps than f_S, permitting shorter rollout lengths: in our running example, one model step corresponds to 200 internal steps of the classical simulator. See Figure 1a for an illustration of simulation using a learned model. To improve stability for longer rollouts, random Gaussian noise was added to the inputs during training.

To optimize a physical design, we leverage the pipeline shown in Figure 1: (1) transform design parameters into an initial scene, (2) simulate the scene using f_M or f_S, (3) evaluate how well the simulation achieves the desired behavior, and (4) adjust the design parameters accordingly.

Design parameters To produce the initial state G^{t_0}, we introduce a differentiable design function $f_D : \Phi \times \mathcal{A} \to \mathcal{G}$ which maps design parameters $\phi \in \Phi$ and other initial conditions $\alpha \in \mathcal{A}$ to an initial state: $G^{t_0} = f_D(\phi, \alpha)$. In Figure 1, ϕ is the 2D height field of the mesh, while α is the non-controllable objects in the scene like the initial position of the fluid.

Maximizing reward The reward function $f_R : \mathcal{G} \times \Theta_R \to \mathbb{R}$ maps the final state of a length-K trajectory (G^{t_K} or $\tilde{G}^{t_K}$) and parameters $\theta_R \in \Theta_R$ to a scalar value. In our running example, the reward function is defined as the Gaussian likelihood of each fluid particle under the closest "pool", averaged across particles (Figure 1a).

We define the objective under the ground-truth simulator f_S as $J_S(\phi) := f_R(f_S^{(K)}(f_D(\phi, \alpha)); \theta_R)$, where $f_S^{(K)}$ indicates K applications of the simulator. We want to find the design parameters that maximize J_S, i.e. $\phi^* = \operatorname{argmax}_\phi J_S(\phi)$. We can approximate this optimization using a *learned* simulation model by maximizing $J_M(\phi) := f_R(f_M^{(K)}(f_D(\phi, \alpha)); \theta_R)$. Crucially, multiple reward functions can be optimized with the same model f_M.

Optimizers Optimal design parameters ϕ^* can be found using any generic optimization technique. Given the differentiability of the learned simulator f_M, we are particularly interested in evaluating gradient-based optimization, which requires fewer function evaluations and scales better to large design spaces than sampling-based techniques [10]. We use the Adam optimizer [44] to find ϕ^* by computing the gradient $\nabla_\phi J_M(\phi)$. This involves backpropagating gradients through the reward function f_R, length-K rollout produced by $f_M^{(K)}$, and design function f_D.

As a baseline, we consider the cross-entropy method (CEM) [59], a gradient-free sampling-based technique that is popular in model-based control [20, 71]. CEM can be used with any simulator and works by sampling a population of candidates for ϕ and evolving them to maximize the reward. However, CEM requires multiple evaluations of f_M or f_S per optimizer iteration (depending on the population size), whereas Adam requires only a single evaluation per iteration.

Across our design tasks (Section 4) we compared: gradient descent with Adam on the learned simulator model (**GD-M**), CEM with the learned simulator model (**CEM-M**), and CEM with the ground-truth simulator (**CEM-S**). We compare to further optimizer baselines, such as Bayesian Optimization [66] and CMA-ES [37] in the Appendix (Section D.2). In all tasks, f_S is non-differentiable, preventing a comparison to GD-S. However, in the special case of AIRFOIL, we compare to **DAFoam** [38], a specialized solver which computes gradients with the adjoint method.

Evaluation Unless otherwise noted, we always evaluate the quality of an optimized design ϕ^* using the ground-truth objective $J_S(\phi^*)$, regardless of whether ϕ^* was found using the learned model f_M (as in CEM-M and GD-M) or the ground-truth simulator f_S (as in CEM-S). We also use rollouts from f_S to produce visualizations in the figures.

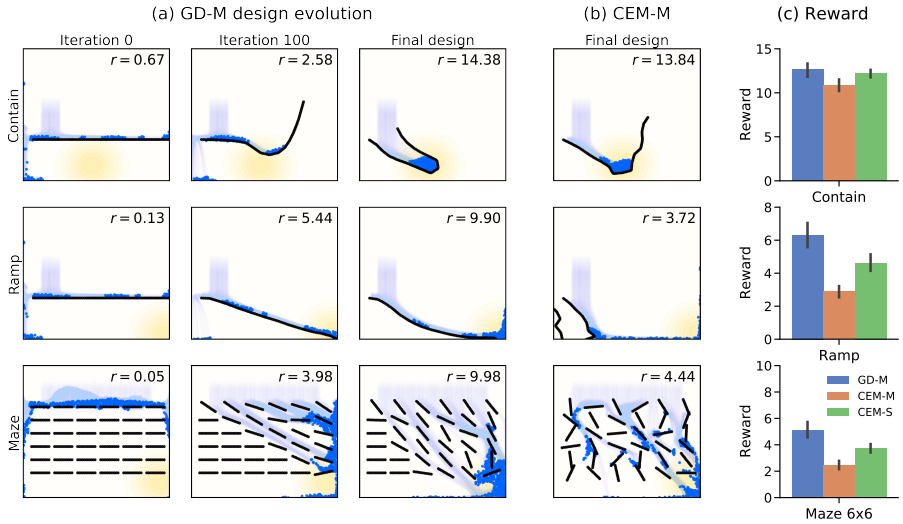

Figure 2: 2D FLUID TOOLS results. The state spaces consist of 10^2–10^3 particles and the design spaces of 16–36 parameters. (a) Evolution of designs found by GD-M during optimization for each 2D FLUID TOOLS task. Visualizations correspond to simulations of the designs under f_S. The design is shown in black, fluid particles in blue, and Gaussian reward in yellow. The transparent particles show the location of fluid for $t < t_K$, and the solid particles show the location of fluid at the final frame $(K = 150)$. r denotes reward for the current design. (b) Final designs found by CEM-M. (c) Mean reward over 50 reward locations (with bootstrapped 95% confidence intervals) obtained by each optimizer across the 2D FLUID TOOLS tasks. For *Contain* and *Ramp*, results are shown for 16 joints; for *Maze*, results are shown for a 6×6 grid of 36 rotors. Across these tasks, GD-M outperforms both CEM-M and CEM-S.

4 Design Tasks

We formulate a set of design tasks across three different physical domains with high-dimensional state spaces and complex dynamics. Each domain uses a different ground-truth simulator f_S, which is used to evaluate designs and pre-train the learned simulator model f_M in a task-agnostic way. During optimization, all the design tasks on each domain share the same pre-trained model.

2D FLUID TOOLS Inspired by existing 2D physical reasoning benchmarks [1, 5], these tasks involve creating one or more 2D "tool" shapes to direct fluid into a particular goal region (Figure 2, Section C.1). We consider three tasks with 4–48 design parameters where the goal is to guide the fluid such that each particle comes as close as possible to the center of a randomly-sampled yellow reward region. In *Contain*, the joint angles ϕ_{joints} of a multi-segment tool must be optimized to catch the fluid by creating cup- or spoon-like shapes. In *Ramp*, the joint angles ϕ_{joints} of a multi-segment tool must be optimized to guide the fluid to a distant location. In *Maze*, the rotation ϕ_{rot} of multiple tools must be optimized to funnel the fluid to the target location. Fluid dynamics are represented using 10^2–10^3 particles, and unrolled for up to 300 time steps.

The learned model f_M is trained on the 2D WaterRamps dataset released in [61] which uses the solver in [41] to generate trajectories. The dataset contains scenes with 1–4 straight line segments; no curved lines or large numbers of obstacles are shown. Therefore, designs which solve 2D FLUID TOOLS tasks are necessarily out of distribution for the learned model.

3D WATERCOURSE To evaluate higher-dimensional inverse design, we created a "landscaping" task that requires optimizing a 3D surface to guide fluids into different areas of an environment (Figure 3, Section C.2). Specifically, water flows out of a pipe and onto an obstacle parameterized by ϕ_{map}, a 25×25 heightmap (625 design parameters). In *Direction*, ϕ_{map} is optimized to redirect the fluid stream towards a specified direction. In *2 Pools* and *3 Pools*, ϕ_{map} is optimized to split the fluid stream such that particles hitting the floor land as close as possible to one of two or three specified pools. Each task has multiple variants, such as different target directions in *Direction*. The simulation

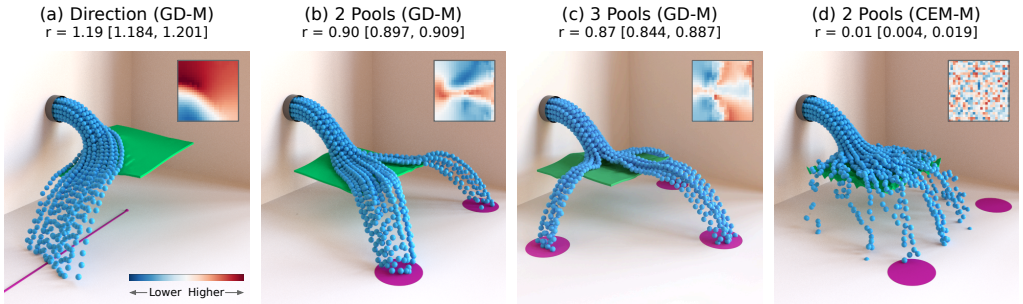

Figure 3: 3D WATERCOURSE results found with GD-M or CEM-M and evaluated with f_S. Simulations use up to 2000 particles and 625 design parameters. The heightmap of a 2D landscape is optimized to redirect the fluid towards the purple targets; birds eye views of heightmaps are shown in the upper right corner of each subplot. In this high-dimensional task domain, GD finds designs with high reward (a-c), while CEM fails to find meaningful solutions for *2 Pools* (d) as well as the other tasks (Figure A.5a). Each subplot reports the mean reward (r) and bootstrapped [lower, upper] 95% confidence intervals for the corresponding optimizer and task (averaged over 10 randomized initial designs for each task variation, see Section C.2).

is unrolled for 50 time steps, and contains up to 2048 particles. The learned model f_M is trained on data generated from the simulator in [8].

AIRFOIL Shape optimization in aerodynamics is one area where gradient-based optimization is routinely applied using traditional simulators [13]. Here, we consider the well-studied task of drag optimization of a 2D airfoil profile (Figure 5, Section C.3). In this task, a wing is defined using a curve on a 2D mesh, which can be deformed using a set of 10 control points ϕ_{ctrl}. The task is formulated as optimizing the wing shape to minimize drag under certain constraints such as constant lift and bounds on the shape to prevent degenerate (e.g. infinitely thin) configurations. Lift and drag coefficients are computed by running an aerodynamics simulation on a 4158 node mesh. The learned model f_M is trained on data generated with the OpenFOAM solver [53].

Model learning and optimization While each domain has a different state space structure and uses a different ground-truth simulator f_S, the learned simulators f_M all share the same architecture (with identical hyperparameters in 2D FLUID TOOLS and 3D WATERCOURSE, and only minor variations of the hyperparameters for AIRFOIL due to it being a steady state simulation; see Appendix B). The models are trained for next-step prediction on task-independent datasets (random perturbations of the design space for AIRFOIL and 3D WATERCOURSE, and an open-source, qualitatively distinct dataset for 2D FLUID TOOLS), and are unrolled for up to 300 time steps during design optimization without further fine-tuning (see Section 3 for details). Importantly, all the pre-trained models are shared across the multiple design tasks for each domain.

5 Results

Our results show that learned simulators can be used to effectively optimize various designs despite significant domain shift and long rollout lengths. The same underlying model architecture is used for each domain, highlighting the generality of learned simulators for design. Examples of designs found with our approach are available at: `https://sites.google.com/view/optimizing-designs`. Performance is always evaluated using the ground-truth simulator f_S (see Section 3). Here we discuss these results, and compare the capabilities of gradient descent with learned simulators over classical simulators and sampling-based optimization techniques.

Overall results We first asked whether a learned simulator combined with gradient descent (GD-M) could produce good-quality designs at all. This approach might fail in various ways: accumulating model error, vanishing or exploding gradients [9], or domain shift [35, 69]. However, as the following results show, GD-M produced high-quality designs across all three domains.

Figure 2a shows qualitative results for 2D FLUID TOOLS (Section 4), where our approach (GD-M) produces intuitive, functional designs to contain (*Contain*), transport (*Ramp*), or funnel (*Maze*) the

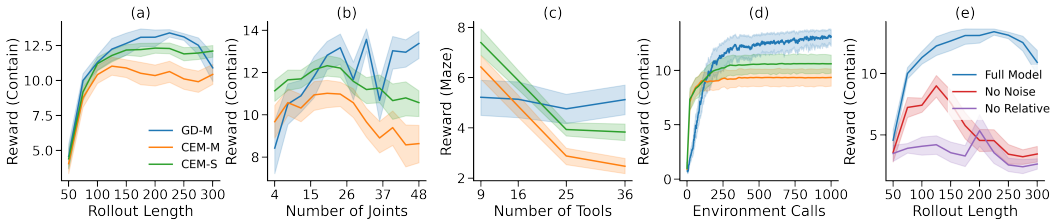

Figure 4: Ablation experiments on the *Contain* and *Maze* tasks. (a) Performance of all optimizers increase with rollout length; GD-M performance starts to deteriorate around step 225. (b) In *Contain*, CEM performance drops when increasing the number of joints above 24, while GD-M remains stable. (c) We observe a similar trend with the number of tools in *Maze*. (d) CEM often gets stuck in sub-optimal solutions early in optimization, while GD-M performance continues to increase. (e) Model choice matters: good performance on GD-M is enabled by key architectural choices of the learned model, such as training noise and translational equivariance.

fluid to a target location. On average, GD-M outperforms CEM-M by 16.1–118.9%, indicating a substantial benefit of gradient-based optimization. GD-M also outperforms CEM-S by 3.9–37.5%, despite using a learned simulator rather than the ground-truth. However, these design spaces are still relatively small (between 16 and 36 dimensions). In 3D WATERCOURSE (Section 4), we substantially increase the dimensionality to a 625-dimensional landscape. Here, GD-M produces robust designs, creating ridges to re-route water in particular directions or valleys to direct water into pools (Figure 3a-c). By comparison, CEM-M cannot solve any of these tasks, with performance 30–85× worse than GD-M. In the Appendix (Section D.2) we demonstrate that these results also hold for other optimizer choices (Bayesian Optimization, CMA-ES).

In AIRFOIL (Figure 5b), GD-M recovers the characteristic S-curve shape for a low-drag airfoil under a small angle of attack, and matches the design obtained with DAFoam, an adjoint aerodynamics solver which computes close-to-optimal designs for this task. The design obtained with DAFoam yields a drag coefficient of 0.01902, while GD-M finds designs with drag between 0.01898–0.01919 depending on ensemble size (see Section 5). Importantly, DAFoam's solver and optimizer are highly specialized for the particular task of airfoil design, while our approach is more general-purpose, requiring only trajectory data for training and a generic gradient-based optimizer.

Model stability & gradient quality We investigated accuracy over long timescales by measuring the effect of rollout length on design quality in 2D FLUID TOOLS (Figure 4a and A.11). Longer rollouts can in principle allow for higher reward in this task as they give the fluid time to settle; however, with learned models, they can also be unstable due to error accumulation [67, 70]. Nevertheless, we find that the learned simulator does not seem to be severely impacted by this problem. Specifically, the quality of designs found by GD-M increases up to 225 steps (Figure 4a), indicating that the learned simulator's accuracy and gradients remain stable for a surprisingly long time. Across the episode lengths evaluated, we find that GD-M outperforms not only CEM-M (by 18.1% on average) but also CEM-S (by 4.4% on average). This indicates that the benefits of a having a learned model that supports better optimization techniques can outweigh the small errors incurred by long rollouts.

The strong performance of GD-M on longer rollout lengths is noteworthy. Gradients tend to degrade when passed through chains of many model evaluations, and as a result, previous work generally only optimizes gradients in small action spaces over just a few time-steps [49].

We hypothesize that a key reason for the success of GD-M is the addition of noise in training the model f_M, which promotes stability on the forward pass and smoother gradients. We test this by considering a model variant trained without noise (Figure 4e, red). We observe much worse design quality compared to the full model, particularly for longer rollouts.

Generalization Deep networks often struggle to generalize far from their training data [31]. This poses a problem for design: to produce in-distribution training data, we would already need to know what good designs look like, thus defeating the aim of wanting to find *new* designs. However, we find that the GNN-based simulators generalize well in the setting studied here. As noted in Section 4, the learned simulator for 2D FLUID TOOLS was trained on a pre-existing, highly simplified dataset where only one to four straight line segments interact with a fluid (Appendix C). However, it still can

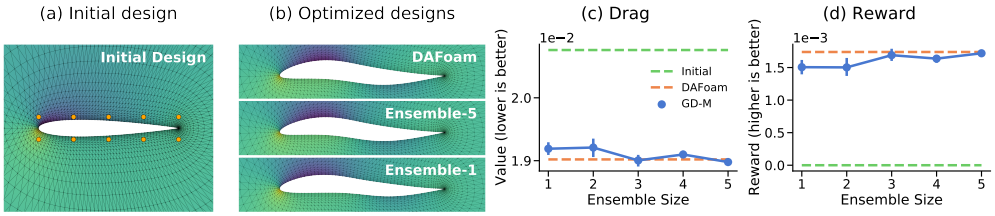

Figure 5: AIRFOIL results. (a) An initial airfoil design is warped by moving 10 control points (orange dots). The physics model simulates the resulting aerodynamics on a 4158 node mesh, based on which lift and drag are computed. (b) For the task of finding a minimum-drag configuration under constant lift constraint, gradient-based learned design is able to find similar designs to specialized solver DAFoam, both using single models and ensembles. (c-d) Larger ensemble sizes of 3–5 achieve a close quantitative match to DAFoam for both drag and overall reward. Shown are means over 10 randomized initial designs, with bootstrapped 95% confidence intervals.

support the design tasks studied here, which involve highly articulated, curved obstacles (*Contain*, *Ramp*) or a larger number of obstacles (*Maze*), without further finetuning (Figure 2).

This generalization capability emerges due to certain design choices of the simulator model f_M. Specifically, the GNN operates on local neighborhoods of particles, and particle coordinates are represented in a relative manner, which encourages learning local rules that can transfer to very different setups. We tested a model variant which breaks this property by encoding *absolute*, instead of *relative* particles coordinates (Figure 4e, purple). We find that this variant can no longer robustly simulate out-of-domain designs, leading to worse performance at any rollout length.

Using a learned simulator trained in a relatively simple environment has another unexpected advantage. In rare cases, classical simulators suffer from degenerate behavior around certain edge cases. For example, with the classical simulator for 2D FLUID TOOLS, particles can get stuck in between joint segments, especially when there are a large number of parts or joints (Figure A.10). However, since the learned simulator was trained on simpler data where these effects are unobserved, it picks up only on the appropriate collision performance and not the unrealistic edge cases. Thus, the learned simulator produces more plausible rollouts than the classical simulator in these cases, and might therefore be a better candidate for producing designs that would transfer to the real world.

Improving accuracy with ensembles In engineering tasks like AIRFOIL, simulators must be especially accurate, as small differences in the predicted pressure field can cause large errors in lift and drag coefficients. While GD-M (without ensembling) can produce designs close to DAFoam's, we notice a slightly rounder wing front (Figure 5b, bottom) causing a small increase in drag (0.01919 versus 0.01902 in DAFoam).

To further improve performance, we implemented an ensemble of learned simulators trained on separate splits of the training set. Ensembles are a popular choice for training transition models for use in control [20], as they can provide higher quality predictions and are more resistant to delusions—a particularly problematic issue for accuracy-sensitive domains such as airfoil design. During optimization, we make predictions with all models in the ensemble, each trained on a different data split, and average the gradients. As shown in Figure 5b-c, larger ensembles yield designs with significantly lower drag ($\beta = -5.3 \times 10^{-5}$, $p = 0.0003$, where β is a linear regression coefficient) and higher overall reward ($\beta = 5.6 \times 10^{-5}$, $p = 0.0001$), and are able to produce designs very close to the solution found by DAFoam, with a drag coefficient of 0.01898 (size-5 ensemble). Thus, with ensembles, we are able to achieve performant designs with a general-purpose learned simulator, indicating that we can use learned models for design optimization in spaces traditionally reserved for specialized solvers like DAFoam.

Scalability to larger design spaces In larger design spaces, sampling-based optimization procedures quickly become intractable, especially with relatively slow simulators. We hypothesized that gradient descent with fast, learned simulators could overcome this issue, especially as the size of the design space is increased. We therefore compared different optimizers on 2D FLUID TOOLS as a function of the dimensionality of the design space (the number of tool joints in *Contain* or the number of tools in *Maze*) and on the higher-dimensional 3D WATERCOURSE.

For *Contain* (Figure 4b), the performance of GD-M increases with the number of joints as increasingly fine grained solutions are made possible. In contrast, for both CEM-M and CEM-S, design quality deteriorates with more joints as high quality solutions become harder to find with random sampling. In the highest dimensional *Contain* task with 48 tool joints, GD-M outperforms CEM-M by 154.9% and CEM-S by 126.5%. When CEM does find solutions (Figure 2b), they lack global coherence and appear more jagged than solutions found with GD-M. Similarly, for *Maze* (Figure 4c), the performance of GD-M is largely unaffected by the number of tools, while the performance of CEM-M and CEM-S both degrade as the design space grows. For the highest dimensional *Maze* problem with 36 joints, GD-M outperforms CEM-M by 207.4% and CEM-S by 133.7%.

In the 625-dimensional 3D WATERCOURSE task, CEM-M performs 30–85$\times$ worse than GD-M (Figure 3) despite extensive hyperparameter tuning. This trend held across all tasks (Figure A.5a), and even persisted when using fewer control points in the design space (Figure A.7). This is due not only to 3D WATERCOURSE's larger design space, but also because this problem requires a globally coherent solution: modifying small areas independently is unlikely to have much effect on the global movement of the fluid.

Model speed and sample efficiency Learned simulators can provide large speedups over traditional simulators in certain domains by learning to compensate for coarser sub-stepping and making optimal use of hardware acceleration. In AIRFOIL, although we use a very simple GD setup, our approach is able to find very similar designs as DAFoam's specialized optimizer. Moreover, our approach requires only 21s (single model) to 62s (size-5 ensemble) on a single A100 GPU, compared to 1021s for DAFoam run on an 8-core workstation, despite requiring $10\times$ more optimization steps.

In 2D FLUID TOOLS, the ground-truth simulator runs at a similar speed to the learned model [see 61], but is non-differentiable and therefore depends on more expensive gradient-free optimization techniques which require more function evaluations. We use 20–40 function evaluations per optimization step of CEM, compared to a single evaluation with GD (which is about $3\times$ more costly, due to the gradient computation). Thus, GD with a differentiable learned model can be much more efficient than using the ground truth simulator with a sampling-based method.

6 Discussion

We find that state-of-the-art learned, differentiable physics simulators can be used with gradient-based optimization to solve challenging inverse design problems. Across three domains and seven tasks, which involved designing landscapes and tools to control water flows or optimizing the shape of an airfoil, we demonstrated that gradient descent with pre-trained simulators can discover high-quality designs that match or exceed the quality of those found using alternative methods. This approach succeeds in a variety of interesting and surprising ways: it permits gradient backpropagation through complex physical trajectories for hundreds of steps; scales to tasks with large design and state spaces (100s and 1000s of dimensions, respectively); and successfully generates designs which require the learned simulator to generalize far beyond its training data. In the classic aerodynamics problem of airfoil shape optimization, our approach produces a design comparable to that of a specialized solver using only simple, general-purpose strategies like model ensembling.

While our results have exciting implications for inverse design, they also open up possibilities for explaining everyday human behavior like tool invention—a longstanding puzzle in cognitive science [1, 54, 65]. With general-purpose learned simulators, we have the potential not just to create highly specialized tools in engineering domains, but also to model everyday tool creation, such as creating a hook from a pipe cleaner, building a blanket fort, or folding a paper boat.

Our approach has limitations that should be visited in future work. Gradient descent is inappropriate for design spaces with regions of zero gradients, such as in 2D FLUID TOOLS tasks where the fluid may not always make contact with the tool (Figure A.9). Many interesting design tasks also have variably-sized or combinatorial design spaces that cannot easily be optimized with gradient descent, such as computer-aided design (CAD) approaches to 3D modeling. An exciting future direction will be to integrate general-purpose learned simulators with hybrid optimization techniques such as those used in material science and robotics [17, 68]. As learned simulators continue to improve, we could also use them to do even broader cross-domain, multi-physics design. While challenges remain, our results represent a promising step towards faster and more general-purpose inverse design.

References

[1] Kelsey R. Allen, Kevin A. Smith, and Joshua B. Tenenbaum. Rapid trial-and-error learning with simulation supports flexible tool use and physical reasoning. *Proceedings of the National Academy of Sciences*, 2020.

[2] W Kyle Anderson and V Venkatakrishnan. Aerodynamic design optimization on unstructured grids with a continuous adjoint formulation. *Computers & Fluids*, 28(4-5):443–480, 1999.

[3] Christof Angermüller, David Belanger, Andreea Gane, Zelda Mariet, David Dohan, Kevin Murphy, Lucy J. Colwell, and D. Sculley. Population-based black-box optimization for biological sequence design. *CoRR*, abs/2006.03227, 2020. URL `https://arxiv.org/abs/2006.03227`.

[4] Aharon Azulay and Yair Weiss. Why do deep convolutional networks generalize so poorly to small image transformations? *CoRR*, abs/1805.12177, 2018. URL `http://arxiv.org/abs/1805.12177`.

[5] Anton Bakhtin, Laurens van der Maaten, Justin Johnson, Laura Gustafson, and Ross Girshick. Phyre: A new benchmark for physical reasoning. *Advances in Neural Information Processing Systems*, 32:5082–5093, 2019.

[6] Peter W Battaglia, Jessica B Hamrick, Victor Bapst, Alvaro Sanchez-Gonzalez, Vinicius Zambaldi, Mateusz Malinowski, Andrea Tacchetti, David Raposo, Adam Santoro, Ryan Faulkner, et al. Relational inductive biases, deep learning, and graph networks. *arXiv preprint arXiv:1806.01261*, 2018.

[7] Filipe de Avila Belbute-Peres, Thomas D. Economon, and J. Zico Kolter. Combining differentiable PDE solvers and graph neural networks for fluid flow prediction. In *Proceedings of the 37th International Conference on Machine Learning ICML 2020*, 2020.

[8] Jan Bender and Dan Koschier. Divergence-free smoothed particle hydrodynamics. In *Proceedings of the 14th ACM SIGGRAPH/Eurographics symposium on computer animation*, pages 147–155, 2015.

[9] Yoshua Bengio, Patrice Simard, and Paolo Frasconi. Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*, 5(2):157–166, 1994.

[10] Homanga Bharadhwaj, Kevin Xie, and Florian Shkurti. Model-predictive control via cross-entropy and gradient-based optimization. In *Learning for Dynamics and Control*, pages 277–286. PMLR, 2020.

[11] James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. JAX: composable transformations of Python+NumPy programs, 2018.

[12] David H. Brookes, Hahnbeom Park, and Jennifer Listgarten. Conditioning by adaptive sampling for robust design, 2019. URL `https://arxiv.org/abs/1901.10060`.

[13] Howard P Buckley, Beckett Y Zhou, and David W Zingg. Airfoil optimization using practical aerodynamic design requirements. *Journal of Aircraft*, 47(5):1707–1719, 2010.

[14] Keith T Butler, Jarvist M Frost, Jonathan M Skelton, Katrine L Svane, and Aron Walsh. Computational materials design of crystalline solids. *Chemical Society Reviews*, 45(22):6138–6146, 2016.

[15] Eduardo F Camacho and Carlos Bordons Alba. *Model predictive control*. Springer science & business media, 2013.

[16] Adithya Challapalli, Dhrumil Patel, and Gouqiang Li. Inverse machine learning framework for optimizing lightweight metamaterials. *Materials & Design*, 208:109937, 2021.

[17] Chun-Teh Chen and Grace X. Gu. Generative deep neural networks for inverse materials design using backpropagation and active learning. *Advanced Science*, 7(5):1902607, 2020.

[18] Tianqi Chen, Bing Xu, Chiyuan Zhang, and Carlos Guestrin. Training deep nets with sublinear memory cost. *arXiv preprint arXiv:1604.06174*, 2016.

[19] Thomas Christensen, Charlotte Loh, Stjepan Picek, Domagoj Jakobović, Li Jing, Sophie Fisher, Vladimir Ceperic, John D. Joannopoulos, and Marin Soljačić. Predictive and generative machine learning models for photonic crystals. *Nanophotonics*, 9(13):4183–4192, 2020.

[20] Kurtland Chua, Roberto Calandra, Rowan McAllister, and Sergey Levine. Deep reinforcement learning in a handful of trials using probabilistic dynamics models. *arXiv preprint arXiv:1805.12114*, 2018.

[21] Shane Colburn and Arka Majumdar. Inverse design and flexible parameterization of meta-optics using algorithmic differentiation. *Communications Physics*, 4(1):1–11, 2021.

[22] Stelian Coros, Bernhard Thomaszewski, Gioacchino Noris, Shinjiro Sueda, Moira Forberg, Robert W. Sumner, Wojciech Matusik, and Bernd Bickel. Computational design of mechanical characters. *ACM Trans. Graph.*, 32(4), jul 2013.

[23] Kyle Cranmer, Johann Brehmer, and Gilles Louppe. The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences*, 117(48):30055–30062, 2020.

[24] Yabo Dan, Yong Zhao, Xiang Li, Shaobo Li, Ming Hu, and Jianjun Hu. Generative adversarial networks (GAN) based efficient sampling of chemical composition space for inverse design of inorganic materials. *Nature Computational Materials*, 6(1):84, 2020.

[25] Marjolein Dijkstra and Erik Luijten. From predictive modelling to machine learning and reverse engineering of colloidal self-assembly. *Nature Materials*, 20(6):762–773, 2021.

[26] Frederik Ebert, Chelsea Finn, Sudeep Dasari, Annie Xie, Alex Lee, and Sergey Levine. Visual foresight: Model-based deep reinforcement learning for vision-based robotic control. *arXiv preprint arXiv:1812.00568*, 2018.

[27] Richard Eppler. *Airfoil design and data*. Springer Science & Business Media, 2012.

[28] Antonio Elia Forte, Paul Z. Hanakata, Lishuai Jin, Emilia Zari, Ahmad Zareei, Matheus C. Fernandes, Laura Sumner, Jonathan Alvarez, and Katia Bertoldi. Inverse design of inflatable soft membranes through machine learning. *Advanced Functional Materials*, page 2111610, 2022.

[29] C Daniel Freeman, Erik Frey, Anton Raichuk, Sertan Girgin, Igor Mordatch, and Olivier Bachem. Brax-a differentiable physics engine for large scale rigid body simulation, 2021.

[30] Justin Fu and Sergey Levine. Offline model-based optimization via normalized maximum likelihood estimation. *CoRR*, abs/2102.07970, 2021. URL `https://arxiv.org/abs/2102.07970`.

[31] Robert Geirhos, Carlos R Medina Temme, Jonas Rauber, Heiko H Schütt, Matthias Bethge, and Felix A Wichmann. Generalisation in humans and deep neural networks. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, pages 7549–7561, 2018.

[32] Justin Gilmer, Samuel S Schoenholz, Patrick F Riley, Oriol Vinyals, and George E Dahl. Neural message passing for quantum chemistry. In *International conference on machine learning*, pages 1263–1272. PMLR, 2017.

[33] Rafael Gómez-Bombarelli, Jennifer N. Wei, David Duvenaud, JoséMiguel Hernández-Lobato, Benjamín Sánchez-Lengeling, Dennis Sheberla, Jorge Aguilera-Iparraguirre, Timothy D. Hirzel, Ryan P. Adams, and Alán Aspuru-Guzik. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Central Science*, 4(2):268–276, 02 2018.

[34] Agrim Gupta, Silvio Savarese, Surya Ganguli, and Li Fei-Fei. Embodied intelligence via learning and evolution. *Nature Communications*, 12(1):5721, Oct 2021.

[35] Jessica B Hamrick, Abram L Friesen, Feryal Behbahani, Arthur Guez, Fabio Viola, Sims Witherspoon, Thomas Anthony, Lars Buesing, Petar Veličković, and Théophane Weber. On the role of planning in model-based deep reinforcement learning. *arXiv preprint arXiv:2011.04021*, 2020.

[36] Paul Z. Hanakata, Ekin D. Cubuk, David K. Campbell, and Harold S. Park. Forward and inverse design of kirigami via supervised autoencoder. *Phys. Rev. Research*, 2:042006, Oct 2020.

[37] Nikolaus Hansen, Pedro Larrañaga, Iñaki Inza, and Endika Bengoetxea. *The CMA Evolution Strategy: A Comparing Review*, pages 75–102. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[38] Ping He, Charles A Mader, Joaquim RRA Martins, and Kevin J Maki. Dafoam: An open-source adjoint framework for multidisciplinary design optimization with openfoam. *AIAA journal*, 58 (3):1304–1319, 2020.

[39] Matteo Hessel, David Budden, Fabio Viola, Mihaela Rosca, Eren Sezener, and Tom Hennigan. Optax: composable gradient transformation and optimisation, in JAX!, 2020.

[40] Stephan Hoyer, Jascha Sohl-Dickstein, and Sam Greydanus. Neural reparameterization improves structural optimization. *CoRR*, abs/1909.04240, 2019.

[41] Yuanming Hu, Yu Fang, Ziheng Ge, Ziyin Qu, Yixin Zhu, Andre Pradhana, and Chenfanfu Jiang. A moving least squares material point method with displacement discontinuity and two-way rigid body coupling. *ACM Transactions on Graphics (TOG)*, 37(4):1–14, 2018.

[42] Yuanming Hu, Luke Anderson, Tzu-Mao Li, Qi Sun, Nathan Carr, Jonathan Ragan-Kelley, and Frédo Durand. Difftaichi: Differentiable programming for physical simulation. *arXiv preprint arXiv:1910.00935*, 2019.

[43] Michael Janner, Justin Fu, Marvin Zhang, and Sergey Levine. When to trust your model: Model-based policy optimization. In *Advances in Neural Information Processing Systems*, 2019.

[44] Diederick P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.

[45] Aviral Kumar, Xue Bin Peng, and Sergey Levine. Reward-conditioned policies. *CoRR*, abs/1912.13465, 2019. URL `http://arxiv.org/abs/1912.13465`.

[46] Siddhant Kumar, Stephanie Tan, Li Zheng, and Dennis M. Kochmann. Inverse-designed spinodoid metamaterials. *Nature Computational Materials*, 6(1):73, 2020.

[47] Charles L Ladson. *Effects of independent variation of Mach and Reynolds numbers on the low-speed aerodynamic characteristics of the NACA 0012 airfoil section*, volume 4074. National Aeronautics and Space Administration, Scientific and Technical . . . , 1988.

[48] Jichao Li, Mengqi Zhang, Joaquim Martins, and Chang Shu. Efficient aerodynamic shape optimization with deep-learning-based geometric filtering. *AIAA Journal*, 58:1–17, 07 2020.

[49] Yunzhu Li, Jiajun Wu, Russ Tedrake, Joshua B. Tenenbaum, and Antonio Torralba. Learning particle dynamics for manipulating rigid bodies, deformable objects, and fluids. In *International Conference on Learning Representations*, 2019.

[50] Zhaocheng Liu, Dayu Zhu, Sean P. Rodrigues, Kyu-Tae Lee, and Wenshan Cai. Generative model for the inverse design of metasurfaces. *Nano Letters*, 18(10):6570–6576, Sep 2018. ISSN 1530-6992.

[51] Michael Lutter, Leonard Hasenclever, Arunkumar Byravan, Gabriel Dulac-Arnold, Piotr Trochim, Nicolas Heess, Josh Merel, and Yuval Tassa. Learning dynamics models for model predictive agents. *arXiv preprint arXiv:2109.14311*, 2021.

[52] Sean Molesky, Zin Lin, Alexander Y Piggott, Weiliang Jin, Jelena Vucković, and Alejandro W Rodriguez. Inverse design in nanophotonics. *Nature Photonics*, 12(11):659–670, 2018.

[53] OpenCFD Ltd. OpenFOAM CFD solver. www.openfoam.org, 2021.

[54] François Osiurak and Arnaud Badets. Tool use and affordance: Manipulation-based versus reasoning-based approaches. *Psychological review*, 123(5):534, 2016.

[55] Tobias Pfaff, Meire Fortunato, Alvaro Sanchez-Gonzalez, and Peter Battaglia. Learning mesh-based simulation with graph networks. In *International Conference on Learning Representations*, 2021.

[56] Janeita Reid. Free-form deformation. dafoam.github.io/docs/FFD/main.pdf, 2021.

[57] Simiao Ren, Willie Padilla, and Jordan Malof. Benchmarking deep inverse models over time, and the neural-adjoint method. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 38–48. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper/2020/file/007ff380ee5ac49ffc34442f5c2a2b86-Paper.pdf.

[58] Li Rhie, C Chow. Numerical study of the turbulent flow past an airfoil with trailing edge separation. *AIAA Journal*, 21(11):1525–1532, 1983.

[59] Reuven Y Rubinstein and Dirk P Kroese. The cross-entropy method: A unified approach to monte carlo simulation, randomized optimization and machine learning. *Information Science & Statistics, Springer Verlag, NY*, 2004.

[60] Alvaro Sanchez-Gonzalez, Nicolas Heess, Jost Tobias Springenberg, Josh Merel, Martin Riedmiller, Raia Hadsell, and Peter Battaglia. Graph networks as learnable physics engines for inference and control. In *International Conference on Machine Learning*, pages 4470–4479. PMLR, 2018.

[61] Alvaro Sanchez-Gonzalez, Jonathan Godwin, Tobias Pfaff, Rex Ying, Jure Leskovec, and Peter Battaglia. Learning to simulate complex physics with graph networks. In *International Conference on Machine Learning*, pages 8459–8468. PMLR, 2020.

[62] Connor Schenck and Dieter Fox. Spnets: Differentiable fluid dynamics for deep neural networks. *Conference on Robot Learning (CoRL)*, 2018.

[63] Ney Secco, Gaetan K. W. Kenway, Ping He, Charles A. Mader, and Joaquim R. R. A. Martins. Efficient mesh generation and deformation for aerodynamic shape optimization. *AIAA Journal*, 2021.

[64] Wuxin Sha, Yan Li, Shun Tang, Jie Tian, Yuming Zhao, Yaqing Guo, Weixin Zhang, Xinfang Zhang, Songfeng Lu, Yuan-Cheng Cao, and Shijie Cheng. Machine Learning in polymer informatics. *InfoMat*, 3(4):353–361, 2021.

[65] Robert W Shumaker, Kristina R Walkup, and Benjamin B Beck. *Animal tool behavior: the use and manufacture of tools by animals*. JHU Press, 2011.

[66] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical bayesian optimization of machine learning algorithms. In F. Pereira, C.J. Burges, L. Bottou, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc., 2012. URL https://proceedings.neurips.cc/paper/2012/file/05311655a15b75fab86956663e1819cd-Paper.pdf.

[67] Erik Talvitie. Model regularization for stable sample rollouts. In *UAI*, pages 780–789, 2014.

[68] Marc A Toussaint, Kelsey Rebecca Allen, Kevin A Smith, and Joshua B Tenenbaum. Differentiable physics and stable modes for tool-use and manipulation planning. In *Proceedings of the Robotics: Science and Systems, RSS 2018*, 2018.

[69] Brandon Trabucco, Aviral Kumar, Xinyang Geng, and Sergey Levine. Conservative objective models for effective offline model-based optimization. In *International Conference on Machine Learning*, pages 10358–10368. PMLR, 2021.

[70] Arun Venkatraman, Martial Hebert, and J Andrew Bagnell. Improving multi-step prediction of learned time series models. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.

[71] Tingwu Wang, Xuchan Bao, Ignasi Clavera, Jerrick Hoang, Yeming Wen, Eric Langlois, Shunshi Zhang, Guodong Zhang, Pieter Abbeel, and Jimmy Ba. Benchmarking model-based reinforcement learning. *arXiv preprint arXiv:1907.02057*, 2019.

[72] Jie Xu, Tao Chen, Lara Zlokapa, Wojciech Matusik, Shinjiro Sueda, and Pulkit Agrawal. An end-to-end differentiable framework for contact-aware robot design. *Robotics: Science and Systems*, 2021.

[73] Sunwoong Yang, Sanga Lee, and Kwanjung Yee. Inverse design optimization framework via a two-step deep learning approach: application to a wind turbine airfoil. *Engineering with Computers*, 2022. doi: 10.1007/s00366-022-01617-6. URL `https://doi.org/10.1007/s00366-022-01617-6`.

[74] Sihyun Yu, Sungsoo Ahn, Le Song, and Jinwoo Shin. Roma: Robust model adaptation for offline model-based optimization. *Advances in Neural Information Processing Systems*, 34, 2021.

[75] Li Zheng, Siddhant Kumar, and Dennis M. Kochmann. Data-driven topology optimization of spinodoid metamaterials with seamlessly tunable anisotropy. *Computer Methods in Applied Mechanics and Engineering*, 383:113894, 2021.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes] See e.g. failure cases in Section D.3.1 and Discussion.
   (c) Did you discuss any potential negative societal impacts of your work? [No] We didn't identify anything specific to this paper, the same considerations as for most work in ML apply (e.g. cost/impact of compute).
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [N/A]
   (b) Did you include complete proofs of all theoretical results? [N/A]
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] The core model we use (MeshGraphNets) has an open-source implementation available, as do the solvers we used to for the task domains (cited in the paper). Detailed description of the task setup can be found in Appendix C.
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes] See Appendix A, Appendix B.
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes]
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [No] We didn't calculate the total amount; but model training was done on typical hardware (up to 4 TPUv3/v100 GPUs, see Appendix B).
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes] We didn't use any data, but cite the model, and solvers used for the task domains.
   (b) Did you mention the license of the assets? [No] All of them are available under open-source licenses.

(c) Did you include any new assets either in the supplemental material or as a URL? [No]

(d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]

(e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]

5. If you used crowdsourcing or conducted research with human subjects...

(a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]

(b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]

(c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]