Grasp'D: Differentiable Contact-rich Grasp Synthesis for Multi-fingered Hands

Dylan Turpin, Liquan Wang, Eric Heiden, Yun-Chun Chen,
Miles Macklin, Stavros Tsogkas, Sven Dickinson, Animesh Garg

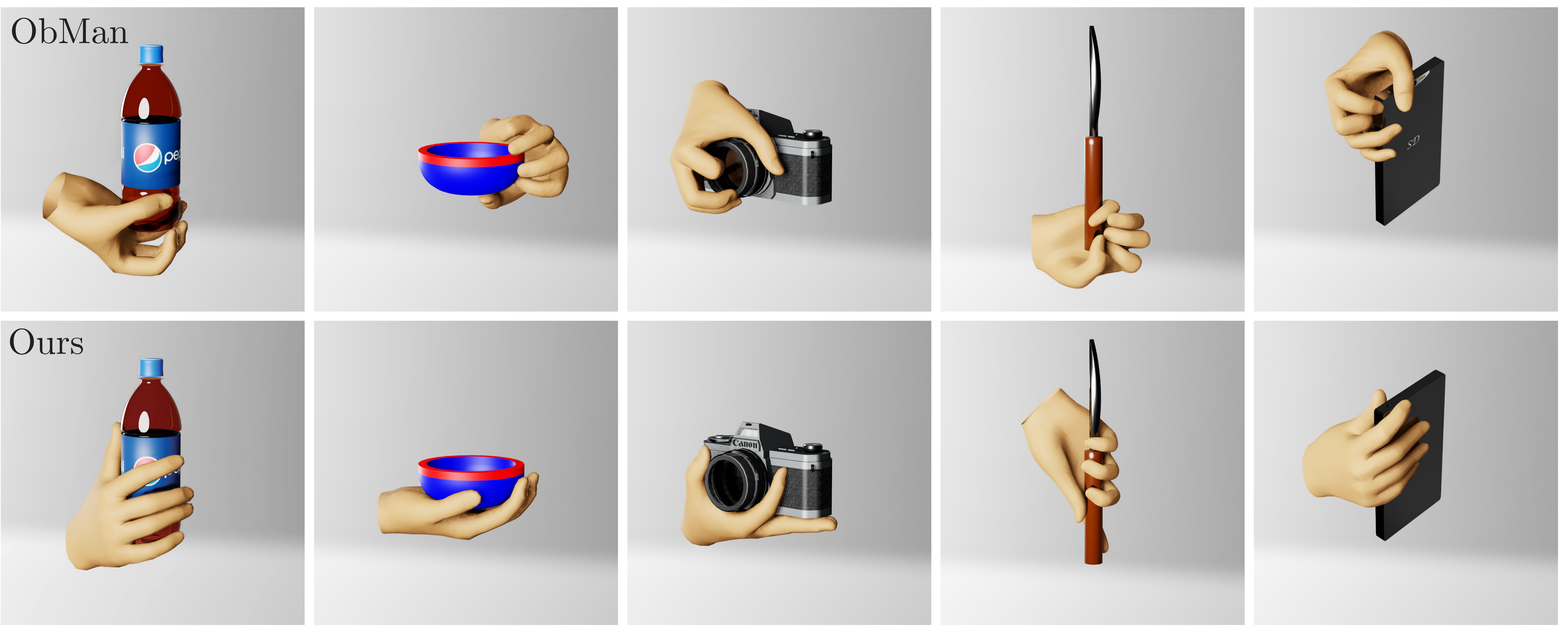
UNIVERSITY OF TORONTO · VECTOR INSTITUTE · NVIDIA · SAMSUNG AI CENTRE

https://graspd-eccv22.github.io/

Problem definition: Given an object mesh, generate a stable grasp as a wrist translation and rotation plus joint angles.

Motivation

Find better grasps faster by leveraging differentiable simulation.



Prior work synthesizes grasps by *blackbox optimization* **(sample intensive)** over *analytic metrics* **(cheap, lower fidelity)**.

Clever *assumptions* **(eigengrasps, preset contacts)** reduce the search space, but result in mainly *fingertip grasps* **(low contact, brittle)**.

We generate grasps by *gradient-based optimization* **(sample efficient)** over a *simulation-based metric* **(expensive, higher fidelity)**.

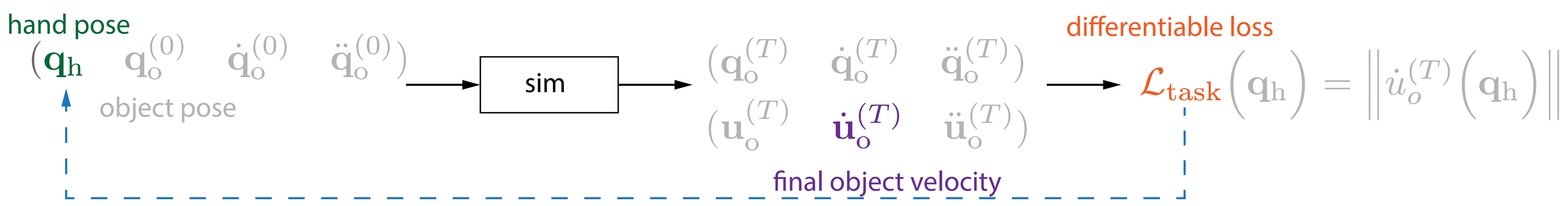
Without assumptions, we discover *contact-rich grasps* **(high contact, stable)** in the full search space.

Challenges and Method

The naive approach to differentiable simulation fails.

1) *Contact sparsity* ⟶ *Leaky gradient*

Contact is sparse! Of all possible hand-object contacts only a few are active at a time. Infinitesimal perturbation of hand pose won't create new contacts, so gradient can vanish.



$$\left(\mathbf{q}_h \; \mathbf{q}_o^{(0)} \; \dot{\mathbf{q}}_o^{(0)} \; \ddot{\mathbf{q}}_o^{(0)}\right) \to \text{sim} \to \begin{pmatrix}\mathbf{q}_o^{(T)} & \dot{\mathbf{q}}_o^{(T)} & \ddot{\mathbf{q}}_o^{(T)} \\ \mathbf{u}_o^{(T)} & \dot{\mathbf{u}}_o^{(T)} & \ddot{\mathbf{u}}_o^{(T)}\end{pmatrix} \to \mathcal{L}_{\text{task}}(\mathbf{q}_h) = \left\|\dot{u}_o^{(T)}(\mathbf{q}_h)\right\|$$

$$\frac{\partial \mathcal{L}_{\text{task}}}{\partial \mathbf{q}_h} = \mathbf{0}$$

So we allow gradient to *leak* through force gradient for inactive contacts.

Proper gradient

$$\frac{\partial \|\mathbf{f}_n\|}{\partial \mathbf{q}} = \begin{cases} k_n \frac{\partial \phi}{\partial \mathbf{q}} & \text{if } \phi(\mathbf{x}) < 0 \\ 0 & \text{otherwise} \end{cases}$$

Leaky gradient

$$\frac{\partial \|\mathbf{f}_n\|}{\partial \mathbf{q}} := \begin{cases} k_n \frac{\partial \phi}{\partial \mathbf{q}} & \text{if } \phi(\mathbf{x}) < 0 \\ \alpha k_n \frac{\partial \phi}{\partial \mathbf{q}} & \text{otherwise} \end{cases}$$

Challenges and Method (cont.)

2) *Non-smooth geometry* ⟶ *Dilated SDF*

Optimizing contacts on a sphere is easy.
Normals and their gradient wrt position is smooth.
But most objects are not so smooth.

We smooth the object surface by first colliding against the radius r>0 level-set.

This gives a smoothed, padded surface that we gradually resolve to the true one by taking r to 0 on a linear schedule.

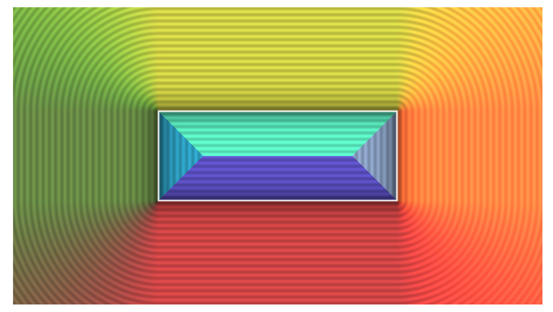


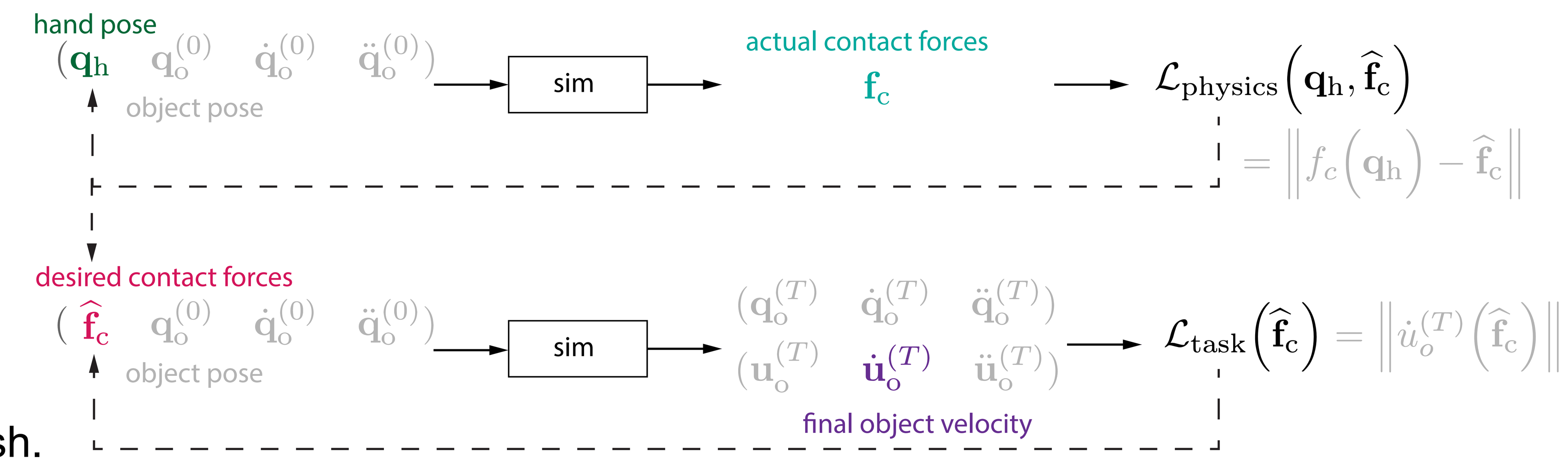
Image source: "2D distance and gradient functions - 2019", Inigo Quilez, https://iquilezles.org/articles/distgradfunctions2d/

3) *Rugged loss landscape* ⟶ *Problem relaxation*

Small changes to hand pose produce large changes in contact forces.
Even with gradient information, optimization is challenging.

We relax the problem by introducing additional variables representing desired contact forces.
This divides the problem into two parts (1, 2).

1 Find hand pose that provides desired contact forces.



$$\left(\mathbf{q}_h \; \mathbf{q}_o^{(0)} \; \dot{\mathbf{q}}_o^{(0)} \; \ddot{\mathbf{q}}_o^{(0)}\right) \to \text{sim} \to \mathbf{f}_c \to \mathcal{L}_{\text{physics}}\left(\mathbf{q}_h, \widehat{\mathbf{f}}_c\right) = \left\|f_c(\mathbf{q}_h) - \widehat{\mathbf{f}}_c\right\|$$

$$\left(\widehat{\mathbf{f}}_c \; \mathbf{q}_o^{(0)} \; \dot{\mathbf{q}}_o^{(0)} \; \ddot{\mathbf{q}}_o^{(0)}\right) \to \text{sim} \to \begin{pmatrix}\mathbf{q}_o^{(T)} & \dot{\mathbf{q}}_o^{(T)} & \ddot{\mathbf{q}}_o^{(T)} \\ \mathbf{u}_o^{(T)} & \dot{\mathbf{u}}_o^{(T)} & \ddot{\mathbf{u}}_o^{(T)}\end{pmatrix} \to \mathcal{L}_{\text{task}}\left(\widehat{\mathbf{f}}_c\right) = \left\|\dot{u}_o^{(T)}\left(\widehat{\mathbf{f}}_c\right)\right\|$$

2 Find desired contact forces that complete the task (as close as possible to current actual contact forces).

Quantitative results

4x more contact, 4x lower displacement

Method	CA ↑	IV ↓	$\frac{CA}{IV}$ ↑	ϵ ↑	Vol ↑	SD ↓
Scale (Unit)	cm^2	cm^3	cm^{-1}	$\times 10^{-1}$	$\times 10^{1}$	cm
ObMan [40] (top2)	9.4	1.28	7.37	4.70	1.36	1.95
ObMan [40] (top5)	7.8	**1.05**	7.37	4.52	1.36	2.22
Grasp'D (top2)	**43.0**	5.70	**7.55**	5.01	1.44	**0.59**
Grasp'D (top5)	41.4	5.48	**7.55**	**5.02**	**1.46**	1.04

4x more contact area (CA) leads to greater stability, i.e., about 4x lower simulation displacement (SD) as well as higher analytic (epsilon and Vol) metrics. Higher contact results in greater interpenetration (IV) but we maintain a similiar CA/IV ratio.

Qualitative results

Plausible, contact-rich, conformal grasps.

Grasps improve as optimization continues.

Grasps conform well to object surface geometry.

Optimizing in the MANO hand PCA weight space gives plausible poses.

For both human and robotic hands.

Replacing the MANO layer with differentiable forward kin. gives a robotic grasp pipeline.

Results vary with initialization, so changing the start pose yields a variety of grasps.

Can also work with multi-view RGBD.

The same pipeline can synthesize grasps from multiview RGBD, by simulating over reconstructed surfaces.

Conclusion

Leveraging differentiable simulation, we find contact-rich, physically plausible grasps often missed by analytic methods.

The sample efficiency provided by gradient-based optimization lets us search the full high-dimensional pose space while spending more computation evaluating each grasp.

Future work should focus on accelerating the pipeline (currently ~5m/grasp or ~20s amortized with 1-gpu parallelism) and on integrating with existing ML pipelines (e.g., integrating a learned adversarial loss or using for vision-based grasping with more advanced reconstruction).