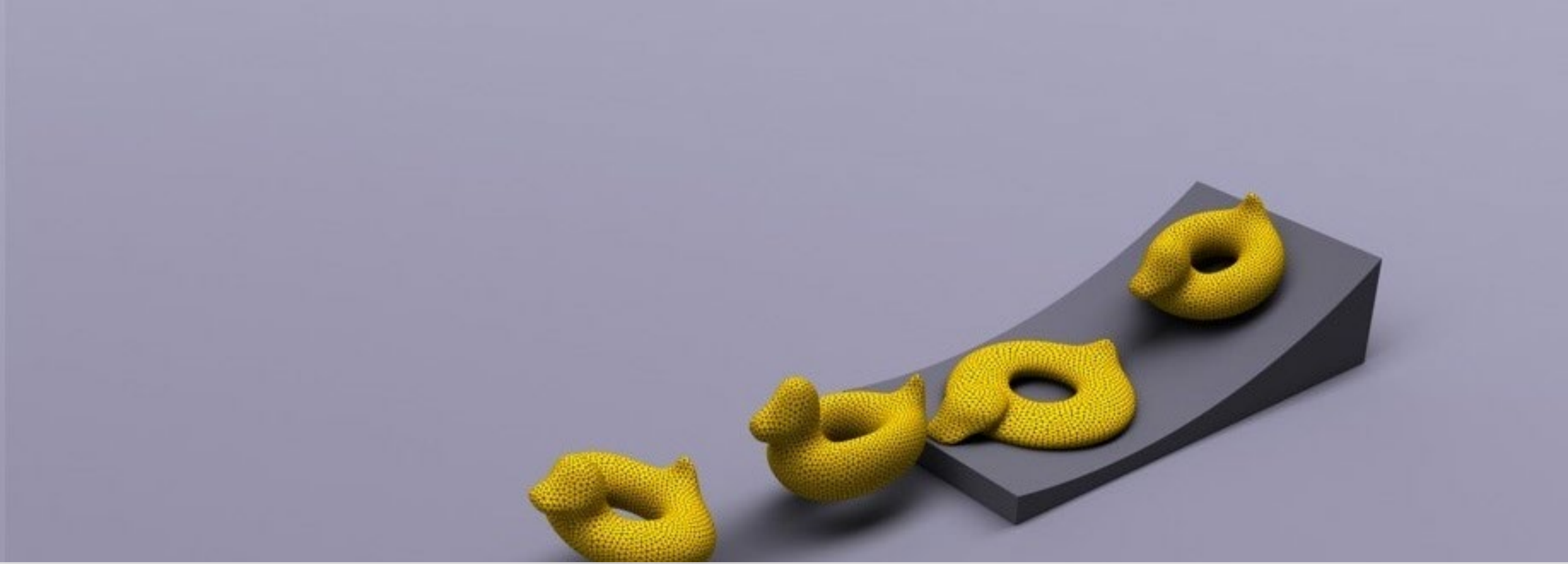


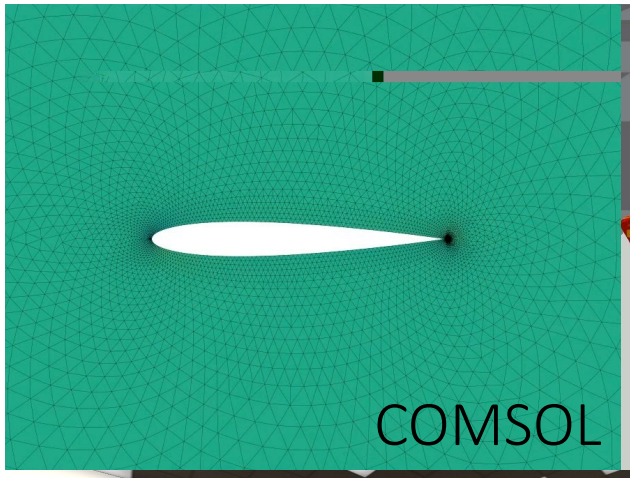
DiffPD: Differentiable Projective Dynamics

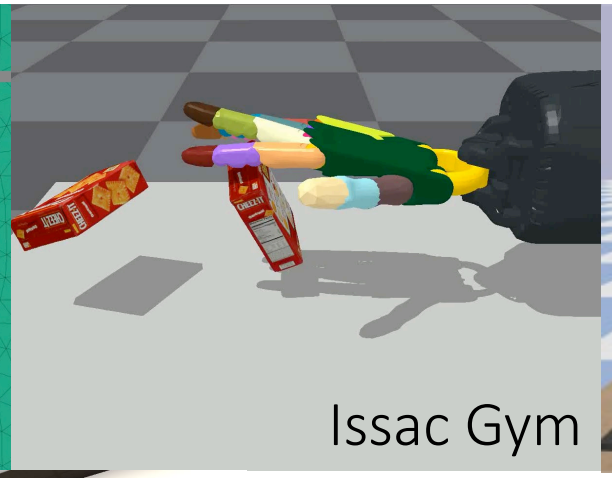
Tao Du, Kui Wu, Pingchuan Ma, Sebastien Wah, Andrew Spielberg,

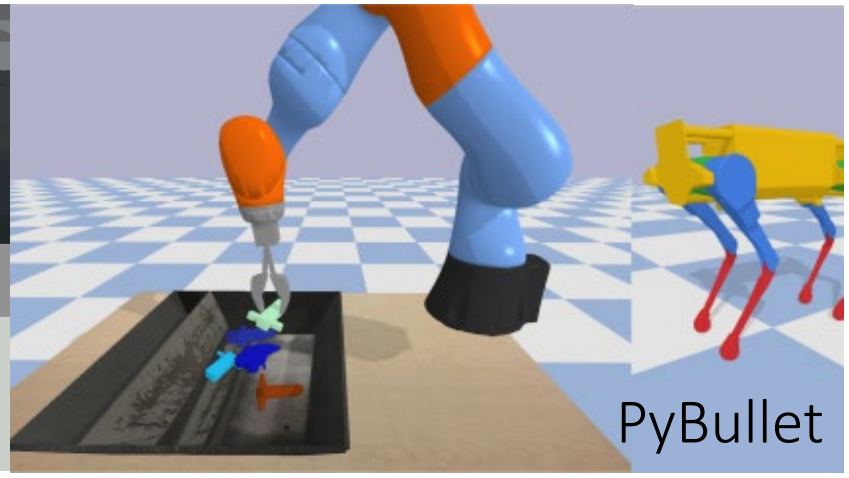
Daniela Rus, Wojciech Matusik

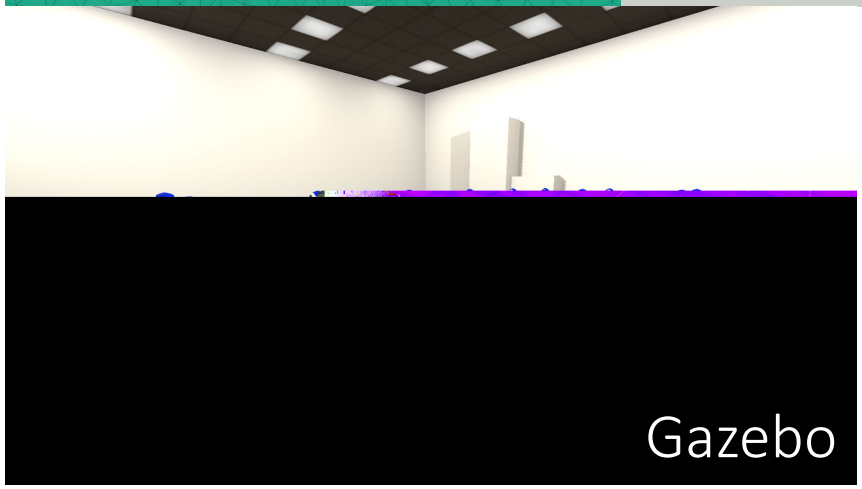
MIT CSAIL

Simulation platforms for learning and robotics









REDMAX

MuJoCo

Method

Background: implicit time integration

Consider the ‴Ω-th step of simulation with timestep ″Ω.

Input: nodal position ȯ$_{ш}$ and velocity ˜I$_{ш}$.

Output: new nodal position ȯ$_{штп}$.

$$ȯ_{штп} = ȯ_{ш} + ″Ω˜I_{штп}$$

$$˜I_{штп} = ˜I_{ш} + ″Ω‘E\,ˁп[ср\ ‘O(ȯ_{штп}) + ˜Hдгдк]$$

Background: implicit time integration

Consider the n-th step of simulation with timestep Δt.

Input: nodal position x_n and velocity v_n.

Output: new nodal position x_{n+1}.

$$x_{n+1} = x_n + \Delta t v_{n+1}$$
$$v_{n+1} = v_n + \Delta t M^{-1}[f_{int}(x_{n+1}) + f_{ext}]$$

mass matrix

Background: implicit time integration

Consider the n-th step of simulation with timestep Δt.

Input: nodal position x_n and velocity v_n.

Output: new nodal position x_{n+1}.

$$x_{n+1} = x_n + \Delta t v_{n+1}$$
$$v_{n+1} = v_n + \Delta t M^{-1}[-\nabla E(x_{n+1}) + f_{ext}]$$

mass matrix

internal force from elastic energy

Background: implicit time integration

Consider the n-th step of simulation with timestep h.

Input: nodal position x_n and velocity v_n.

Output: new nodal position x_{n+1}.

$$x_{n+1} = x_n + h v_{n+1}$$
$$v_{n+1} = v_n + h M^{-1}[-\nabla E(x_{n+1}) + f_{ext}]$$

mass matrix

internal force from elastic energy

external force

Background: implicit time integration

Recast it as a saddle-point problem: find $\text{‘}\Omega(\text{ȯ}_{\text{ﺱﺙپ}}) = \text{ج}$ where

$$\text{‘}\Omega(\text{ȯ}) \ \text{ḧ} \ \frac{1}{2\text{”}\Omega_{\text{پ}}}(\text{ȯ سر ȯ})^{\text{ڑ}} \text{‘E}(\text{ȯ سر ȯ}) + \text{‘O}(\text{ȯ})$$

Background: implicit time integration

Recast it as a saddle-point problem: find $\hat{u}(\dot{o}_{n+1}) = \hat{g}$ where

$$\hat{u}(\dot{o}) \equiv \frac{1}{2\hat{\Omega}\hat{\rho}}(\dot{o} - \dot{o}^{n})^{T}\hat{E}(\dot{o} - \dot{o}^{n}) + \hat{O}(\dot{o})$$

$\dot{o} \equiv \dot{o}_{n} + \hat{\Omega}\tilde{I}_{n} + \hat{\Omega}\hat{\rho}\hat{E}^{-1}\tilde{H}_{n}\hat{g}_{n}$ is independent of $\dot{o}$.

Background: implicit time integration

Newton's method: $\dot{o}^{m+1} = \dot{o}^{m} + \delta\dot{o}^{m}$ where

$$J(\dot{o}^{m})\delta\dot{o}^{m} = -R(\dot{o}^{m})$$

Bottleneck: solving the matrix $J(\dot{o}^{m})$:

$$J(\dot{o}^{m}) = \frac{1}{\Delta t}M + K(\dot{o}^{m})$$

Background: implicit time integration

Newton's method: $\dot{o}^{\text{ش}\text{ت}\text{پ}} = \dot{o}^{\text{ش}} + \text{ع}\dot{o}^{\text{ش}}$ where

$$\text{پ}''\Omega\left(\dot{o}^{\text{ش}}\right)\text{ع}\dot{o}^{\text{ش}} = \quad ''\Omega\left(\dot{o}^{\text{ش}}\right)$$

Bottleneck: solving the matrix $\text{پ}''\Omega\left(\dot{o}^{\text{ش}}\right)$:

$$\text{پ}''\Omega\left(\dot{o}^{\text{ش}}\right) = \frac{1}{''\Omega\text{پ}}\,'\text{E} + \quad \text{پ}'\text{O}\left(\dot{o}^{\text{ش}}\right)$$

requires recomputation whenever $\dot{o}^{\text{ش}}$ changes!

Background: differentiable simulation

Consider backpropagating loss ℓ in the t-th step of simulation.

$$\frac{\partial \ell}{\partial s} = \frac{\partial \ell}{\partial s_{t+1}} \frac{\partial s_{t+1}}{\partial s}$$

Recall that $f(s_{t+1}) = c$. By differentiating it w.r.t. s we have

$$\frac{\partial s_{t+1}}{\partial s} = \frac{1}{f'} \left[\quad f' f(s_{t+1}) \right]^{T} f'E$$

Background: differentiable simulation

Putting everything together, we have $\frac{\partial \mathcal{L}}{\partial \theta} = \frac{\partial \mathcal{L}}{\partial s'} \frac{\partial s'}{\partial \theta}$ where

$$\frac{\partial s'}{\partial \theta}(\theta_{sim}) = \left(\frac{\partial s'}{\partial \theta} \right)^{T}$$

Background: differentiable simulation

Putting everything together, we have $\frac{\partial L}{\partial x_0} = \frac{\partial x}{\partial x_0} \nabla_x E$ where

$$\frac{\partial x}{\partial x_0}(x_{t+\Delta t}) = \left(\frac{\partial x_t}{\partial x_{t+\Delta t}} \right)^T$$

We see that solving $\nabla_x E$, again, is the bottleneck.

Recap

Forward simulation: $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t)$.

Backpropagation: $\nabla_{\mathbf{x}_t} L = \left(\frac{\partial \mathbf{x}_{t+1}}{\partial \mathbf{x}_t}\right)^T \nabla_{\mathbf{x}_{t+1}} L$.

Insight

Forward and backward computation share the same bottleneck.

Forward simulation: $\mathcal{O}(s)\mathcal{F}s = \mathcal{O}(s).$

Backpropagation: $\mathcal{O}(s')\nabla s = \left(\frac{\partial L}{\partial s}\right)^{T}.$

Insight

Efficient solvers for forward simulation exist.

Efficient forward simulation: $\mathcal{O}(n)$

Backpropagation: $\mathcal{O}(n)$

Insight

Can we borrow them to build efficient backpropagation solver as well?

Efficient forward simulation: $\mathcal{L}(o^{t+1}) = \mathcal{L}(o^t)$.

Efficient backpropagation: $\frac{\partial \mathcal{L}(o^t)}{\partial o^t} = \left(\frac{\partial \mathcal{L}}{\partial o^{t+1}}\right)^T$.

Simulation speedup: Projective Dynamics (PD)

Consider a special class of $\dot{O} = B_r O_r$ where

$$O_r(o) \; \ddot{h} \; \min_{I_{iz}{}^{N} \; \ddot{A}_{iz}} \| \acute{E}_r o \; sr \; I_r \|_{\ddot{E}}^{\ddot{E}}$$

Simulation speedup: Projective Dynamics (PD)

Consider a special class of $\dot{U} = B_r U_r$ where

$$U_r(t+h) \approx \min_{I_{in}} \sum_{i \in n} \| A_i I_r - S_i U_r \|^2$$

energy on each
finite element

Simulation speedup: Projective Dynamics (PD)

Consider a special class of $\dot{U} = B_r U_r$ where

$$U_r(u) \approx \min_{I_{iz} \in \mathbb{R}^N} \| \epsilon_r u \, s_r I_r \|_F^2$$

energy on each finite element

local feature, e.g., deformation gradient

Simulation speedup: Projective Dynamics (PD)

Consider a special class of $\dot{U} = B_r U_r$ where

$$U_r(u) \; \ddot{H} \; \min_{I_{iz}^{N} \; \ddot{A}_{iz}} \| \acute{E}_r u \; s_r \; I_r \| \ddot{I}$$

energy on each finite element

local feature, e.g., deformation gradient

projection of local feature

Simulation speedup: Projective Dynamics (PD)

The saddle-point problem $\mathbf{A}\mathbf{x} = \mathbf{b}$ is now modified accordingly:

$$\min_{\mathbf{x},\{\mathbf{z}_i \in \mathbb{S}_i\}} \; w_M(\mathbf{x}, \{\mathbf{z}_i\})$$

where

$$w_M(\mathbf{x}, \{\mathbf{z}_i\}) \doteq \frac{1}{2}(\mathbf{x}-\mathbf{s})^T \mathbf{M}(\mathbf{x}-\mathbf{s}) + \frac{w_i}{2}\sum_i \|\mathbf{A}_i\mathbf{x} - \mathbf{z}_i\|^2$$

Simulation speedup: Projective Dynamics (PD)

The saddle-point problem $\ddot{\Omega} = \ddot{\jmath}$ is now modified accordingly:

$$\min_{\dot{o},\{\dot{I}_{\dot{\imath}} \in \dot{\Omega}_{\dot{\imath}}\}} \frac{1}{2\ddot{\Omega}_{\dot{\imath}}} (\dot{o} - \dot{s})^{T} E(\dot{o} - \dot{s}) + \frac{\dot{\kappa}}{2} \sum_{\dot{\imath}} \|\dot{R}_{\dot{\imath}} \dot{I}_{\dot{\imath}} - \dot{G}\dot{o}\|_{\dot{\imath}}^{2}$$

Simulation speedup: Projective Dynamics (PD)

The saddle-point problem $\mathbf{A}\mathbf{x} = \mathbf{b}$ is now modified accordingly:

$$\min_{\mathbf{x},\{\mathbf{z}_i\}} \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{x}^T\mathbf{b} + \frac{w}{2}\sum_i \|\mathbf{C}_i\mathbf{x} - \mathbf{z}_i\|^2$$

The global step: fix $\mathbf{z}_i$ and solve $\mathbf{x}$, constant matrix in the quadratic form and can be prefactorized.

Simulation speedup: Projective Dynamics (PD)

The saddle-point problem $\mathbf{A}\mathbf{x} = \mathbf{b}$ is now modified accordingly:

$$\min_{\mathbf{x},\{\mathbf{z}_i \in \mathcal{C}_i\}} \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{x}^T\mathbf{b} + \frac{w}{2}\sum_i \|\mathbf{S}_i\mathbf{x} - \mathbf{z}_i\|_2^2$$

The global step: fix $\mathbf{z}_i$ and solve $\mathbf{x}$, constant matrix in the quadratic form and can be prefactorized.

The local step: fix $\mathbf{x}$ and solve $\mathbf{z}_i$, parallelizable among elements.

Backpropagation speedup: DiffPD

With PD, ؆ becomes

$$\text{؆}(\text{ȯ}) = \frac{1}{\text{؆}}\text{E} + \text{ضحي}_{\text{ر}}\,\text{ڠ}_{\text{ر}}\,\text{س}_{\text{ر}}\,\text{ڠ}_{\text{ر}}\,\text{ضحي}_{\text{ر}}\,\text{ڠ}_{\text{ر}}\frac{\text{I}_{\text{ر}}}{\text{ȯ}}$$

$$\text{ḧ}\ \text{A}\ \text{س}\ \text{ز}\text{A}$$

Backpropagation speedup: DiffPD

With PD, ُ becomes

$$\text{ُ}(\text{ó}) = \frac{1}{\text{ّ}}\,\text{ٔ}\text{E} + \text{ضحي}\ \text{ڑ}\text{غ}\ \text{ر}\ \text{غ}\ \text{ر}\ \text{سر}\ \text{ضحي}\ \text{ر}\ \text{غ}\ \text{ڑ}\ \frac{\text{ر}\,\text{ٔ}\text{I}}{\text{ó}}$$

$$\text{ï}\ \ \text{ٓ}\text{A}\ \text{سر}\ \text{ٓ}\text{A}$$

constant matrix and
source of efficiency

Backpropagation speedup: DiffPD

With PD, ﻱ"Ω becomes

$$ﻱ"Ω(ȯ) = \frac{1}{"Ωﻱ}\,'E + ضحي\ ڑغ\ ر\ سر\ غ\ ڑغ\ ضحي\ ر \frac{ر\ 'I}{ȯ}$$

$$Ḧ\ \ 'A\ سر\ ﻉ"A$$

constant matrix and source of efficiency

residual that can be computed in parallel.

Backpropagation speedup: DiffPD

Recall the bottleneck:

$$\nabla\Omega(x_{t+1})\dot{x} = M\ddot{h}\left(\frac{\partial x}{\partial x_{t+1}}\right)^T$$

With $\nabla\Omega = A^T A$ it becomes $(A^T A)\dot{x} = M\ddot{h}$, or

$$A\dot{x} = M\ddot{h} + A\dot{x}$$

Backpropagation speedup: DiffPD

Recall the bottleneck:

$$\nabla\Omega(x_{step})\dot{x} = \dot{H}\left(\frac{\partial U}{\partial x_{step}}\right)^{T}$$

With $\nabla\Omega = A$ we have $\nabla\Omega\dot{x} = \dot{H}$, or

$$A\dot{x} = \dot{H} + \nabla\Omega\dot{x}$$

The global step: constant A,
already factorized.

Backpropagation speedup: DiffPD

Recall the bottleneck:

$$\nabla\Omega(x_{t+1})\dot{x} = M\ddot{h}\left(\frac{\partial x_{t+1}}{\partial x_t}\right)^T$$

With $\nabla\Omega = A \otimes \mathrm{I}_3$ it becomes $(A \otimes \mathrm{I}_3)\dot{x} = M\ddot{h}$, or

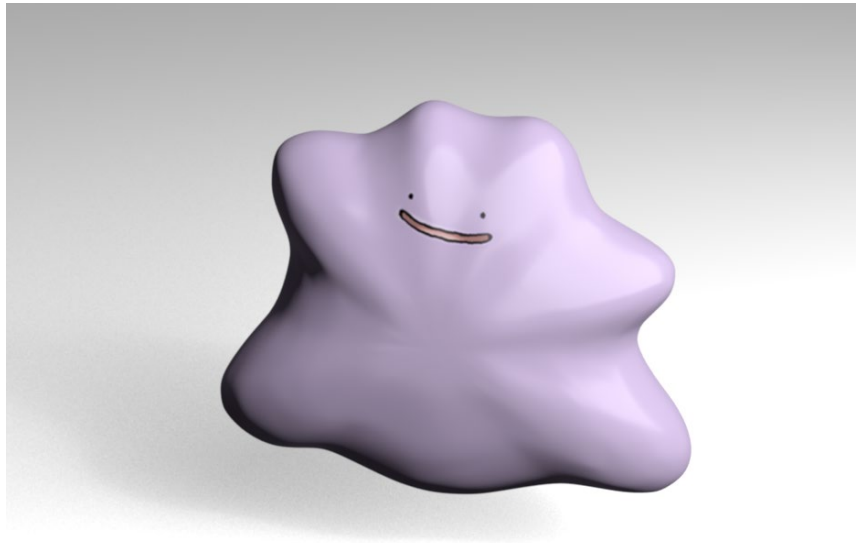
$$A\dot{x} = M\ddot{h} + \mathrm{I}_3\dot{x}$$

The global step: constant A, already factorized.

The local step: parallelizable on elements.

Extension one: quasi-Newton speedup

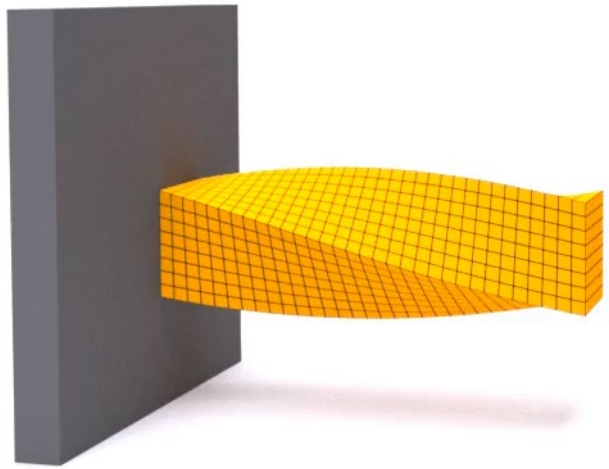
We recall that Liu [2017] proposed a quasi-Newton approach to speed up PD even more. We can transfer it to backpropagation too.

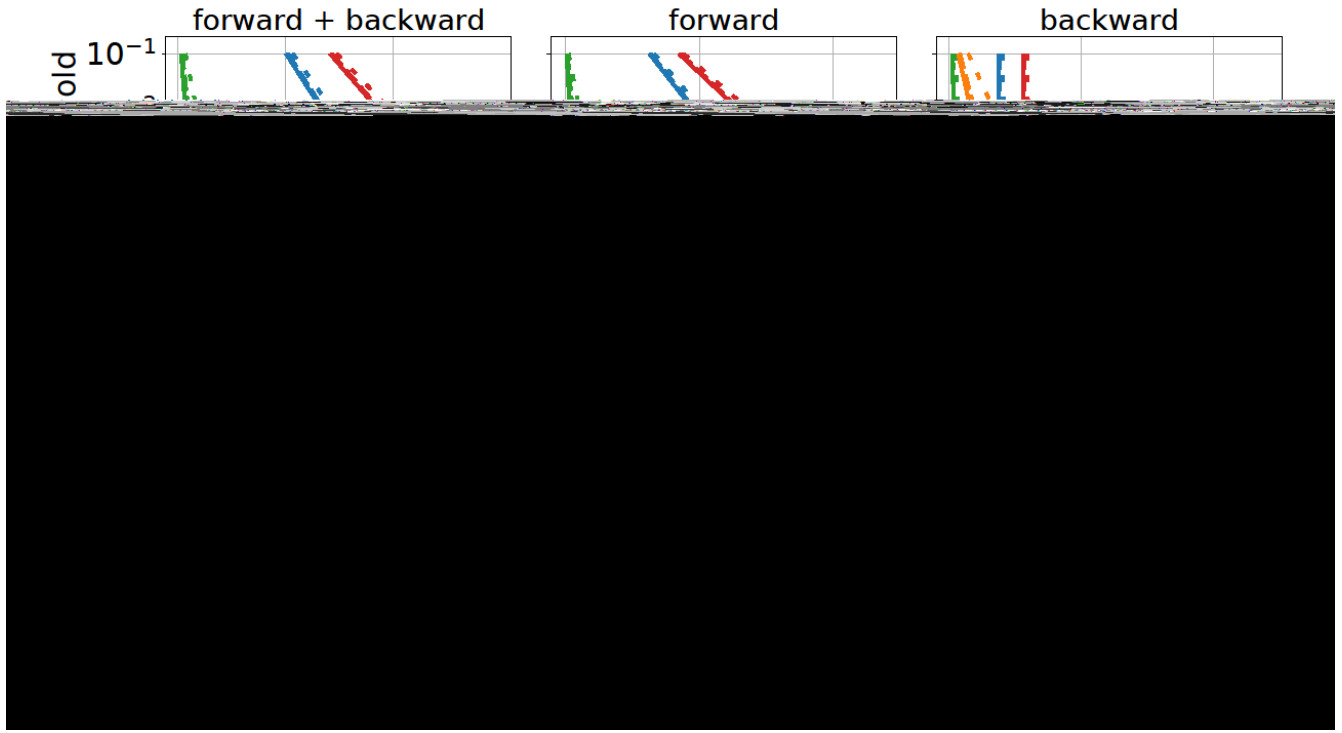


Liu et al. *Quasi-Newton Methods for Real-time Simulation of Hyperelastic Materials*. TOG 2017

Extension one: evaluation

Cantilever (8019 DoFs, 25 steps, 10ms timestep, no contact)





Extension two: boundary conditions

Consider the global step in PD:

$$\text{Ä}\acute{o}_{\text{پتس}} = \text{right-hand side from local step.}$$

Efficient solvers exist for Ä with erased rows and columns:

$$(\text{Ä} + \text{˝I}\tilde{}\text{I}^{\text{ڑ}})^{\text{ڑ}\text{پ}} = \text{Ä}^{\text{ڑ}\text{پ}} \text{سر} \frac{\text{Ä}^{\text{ڑ}\text{پ}}\text{˝I}(\text{Ä}^{\text{ڑ}\text{پ}}\tilde{}\text{I})^{\text{ڑ}}}{1 + \tilde{}\text{I}^{\text{ڑ}}\text{Ä}^{\text{ڑ}\text{پ}}\text{˝I}}$$

Extension two: boundary conditions

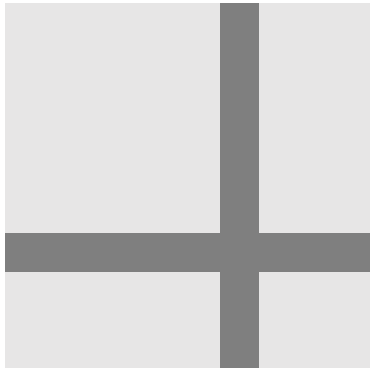
Let’s take a look at its source of efficiency:

$$(\text{“A} + \text{“I}\tilde{\text{I}}^{\text{ڎ}})^{\text{ڟ}} = \text{“A}^{\text{ڟ}} \times \frac{\text{“A}^{\text{ڟ}}\text{“I}(\text{“A}^{\text{ڟ}}\tilde{\text{I}})^{\text{ڎ}}}{1 + \tilde{\text{I}}^{\text{ڎ}}\text{“A}^{\text{ڟ}}\text{“I}}$$

Extension two: boundary conditions

Let’s take a look at its source of efficiency:

$$(\text{“A} + \text{“I}\tilde{\text{I}}^{\text{ڟ}})^{\text{ڜ}} = \text{“A}^{\text{ڜ}} \text{س} \frac{\text{“A}^{\text{ڜ}}\text{“I}(\text{“A}^{\text{ڜ}}\tilde{\text{I}})^{\text{ڟ}}}{1 + \tilde{\text{I}}^{\text{ڟ}}\text{“A}^{\text{ڜ}}\text{“I}}$$



Extension two: boundary conditions

Let’s take a look at its source of efficiency:

$$(‘A + ‘’I\tilde{I}^{\flat})^{\natural} = ‘A^{\natural} \prec \frac{‘A^{\natural}\,‘’I(‘A^{\natural}\,\tilde{I})^{\flat}}{1 + \tilde{I}^{\flat}\,‘A^{\natural}\,‘’I}$$

Extension two: boundary conditions

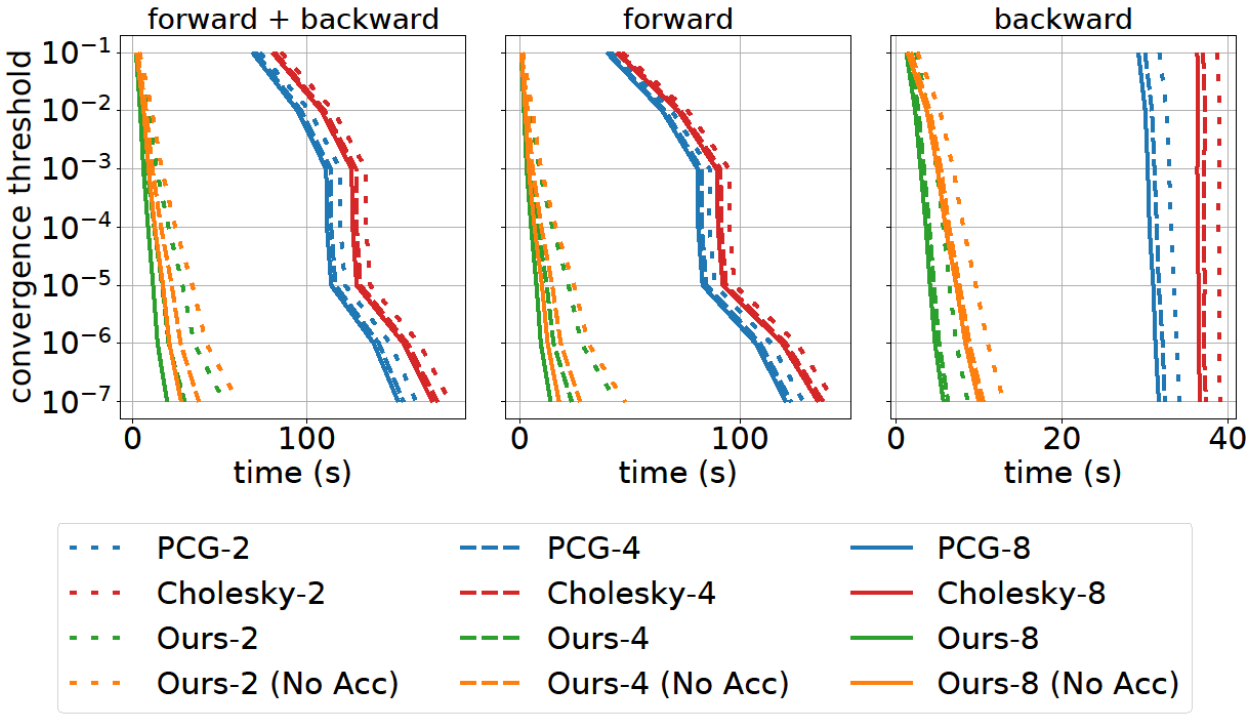
Let’s take a look at its source of efficiency:

$$(\mathbf{A} + \mathbf{I}\tilde{\mathbf{I}}^{T})^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{I}(\mathbf{A}^{-1}\tilde{\mathbf{I}})^{T}}{1 + \tilde{\mathbf{I}}^{T}\mathbf{A}^{-1}\mathbf{I}}$$

It turns out that we can transfer this idea to backpropagation too, which DiffPD used for contact handling.

Extension two: evaluation

Rolling sphere (2469 DoFs, 100 steps, 5ms timestep, **with contact**)



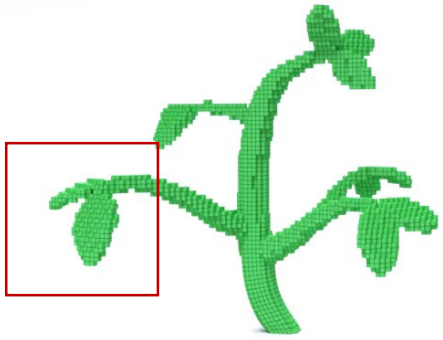
Insight, reiterated

Efficient forward simulation solvers can be transferred to efficient backpropagation solvers!

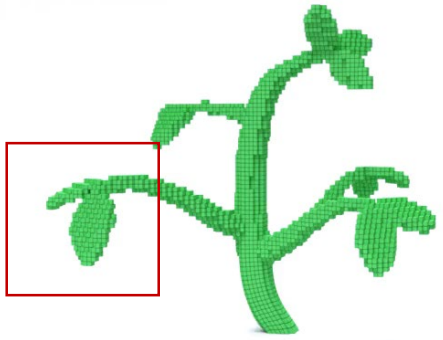
Applications

System identification

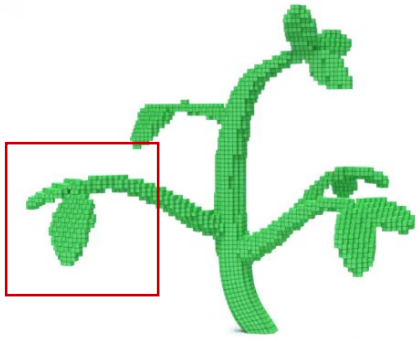
Goal: estimating the material parameters of a plant from its motion.



Input (ground truth)



Initial guess



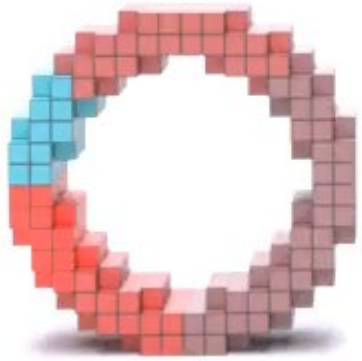
After optimization

Initial state optimization

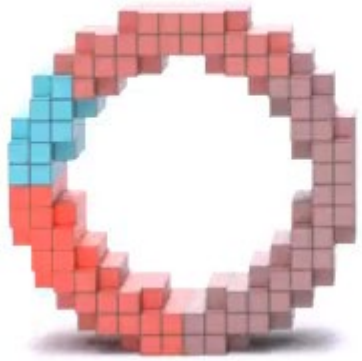
Goal: optimizing time-invariant actuation to reach the target.

Open-loop control

Goal: optimizing actuation to roll forward.



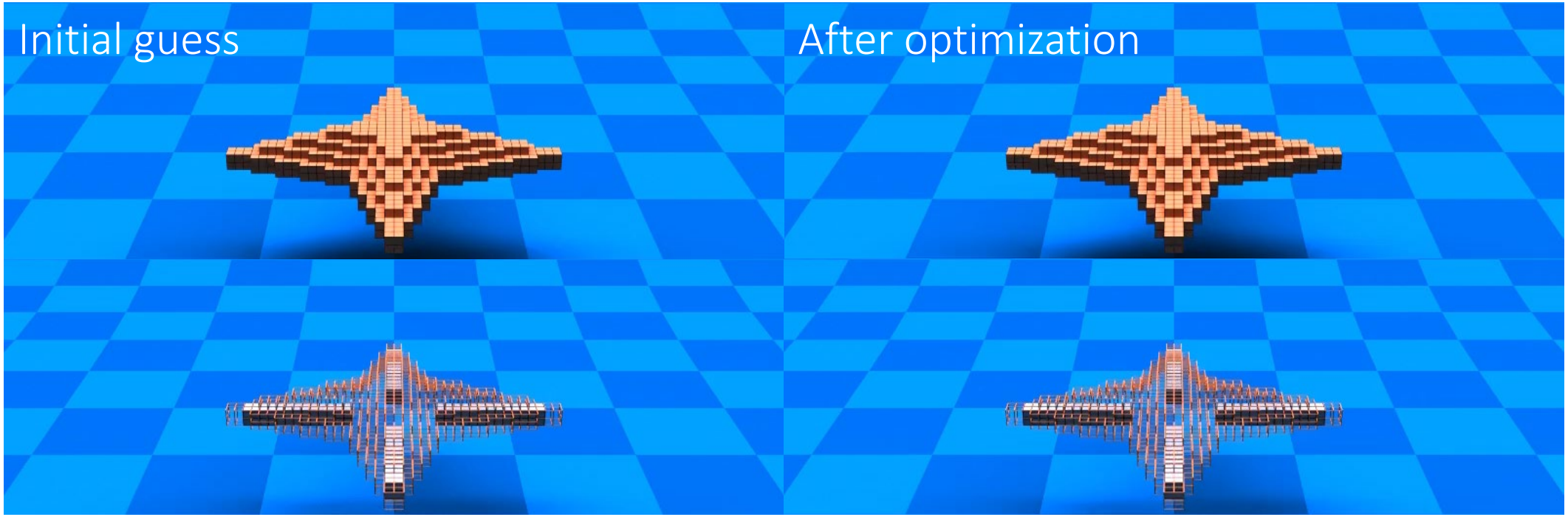
Initial guess



After optimization

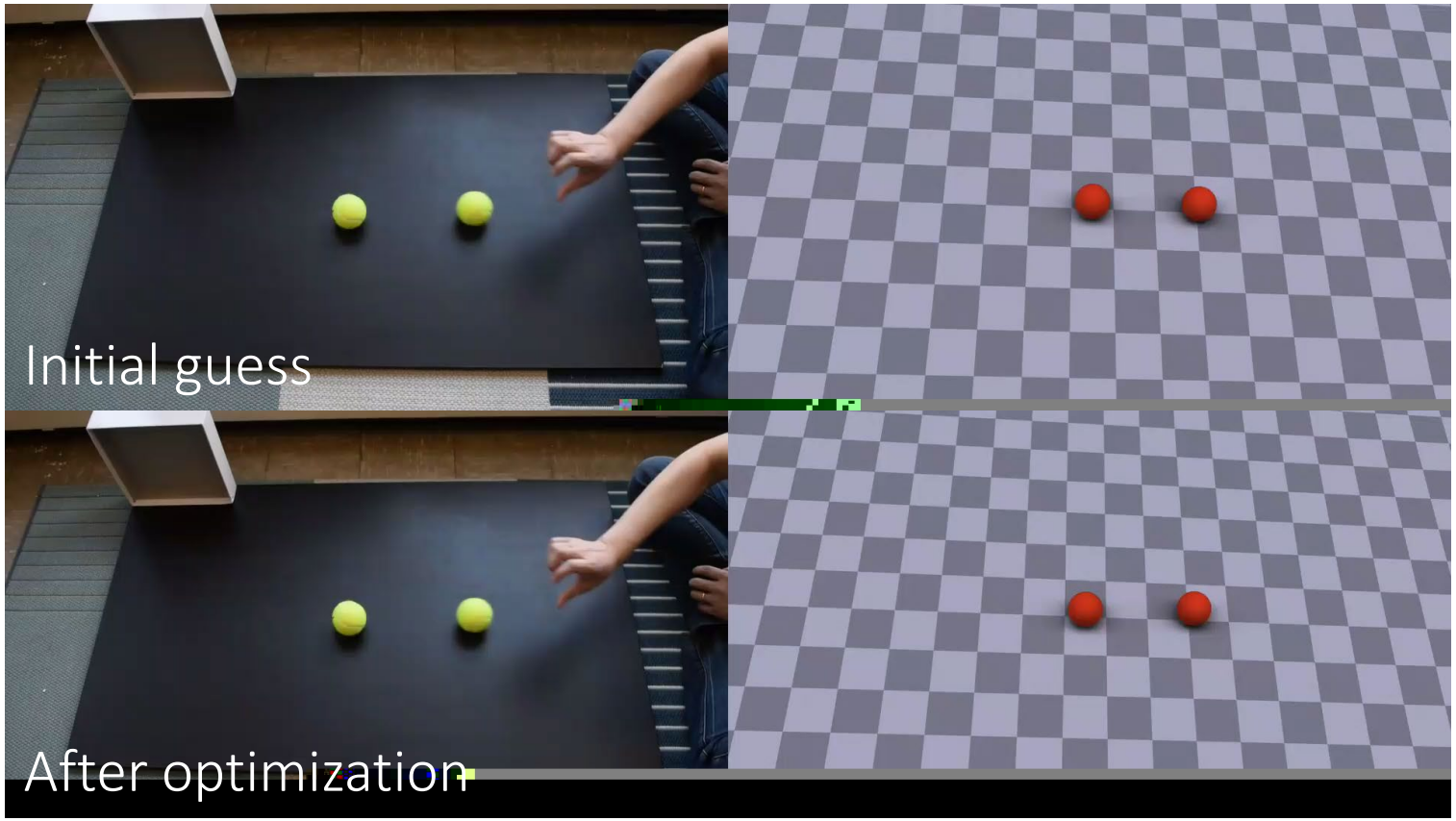
Closed-loop control

Goal: optimizing a neural network controller so that the starfish rises.



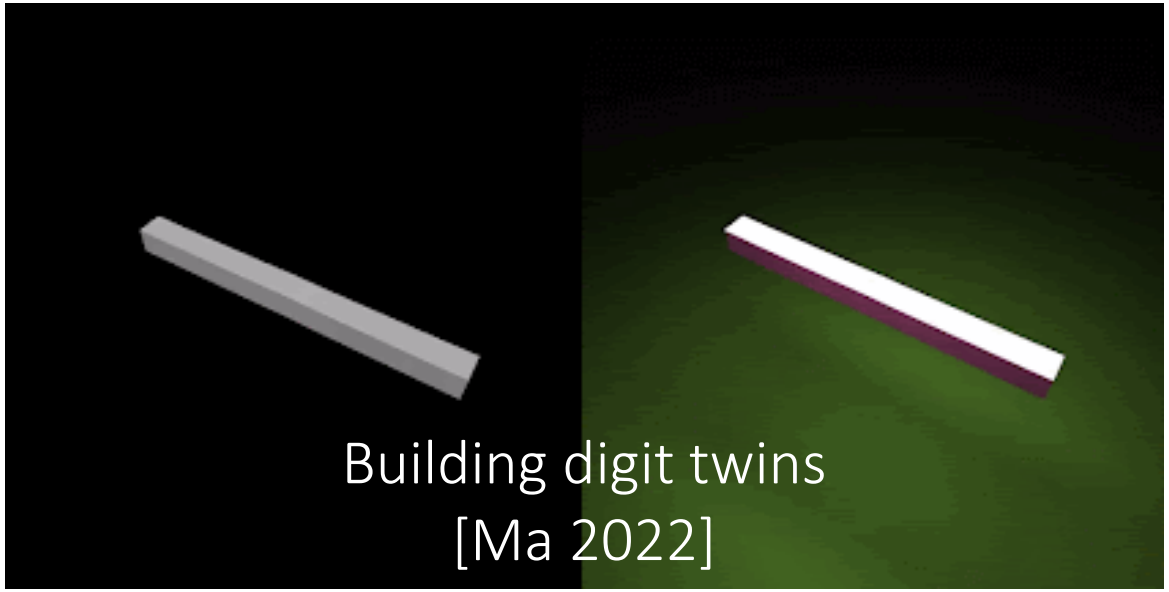
Real-to-sim transfer

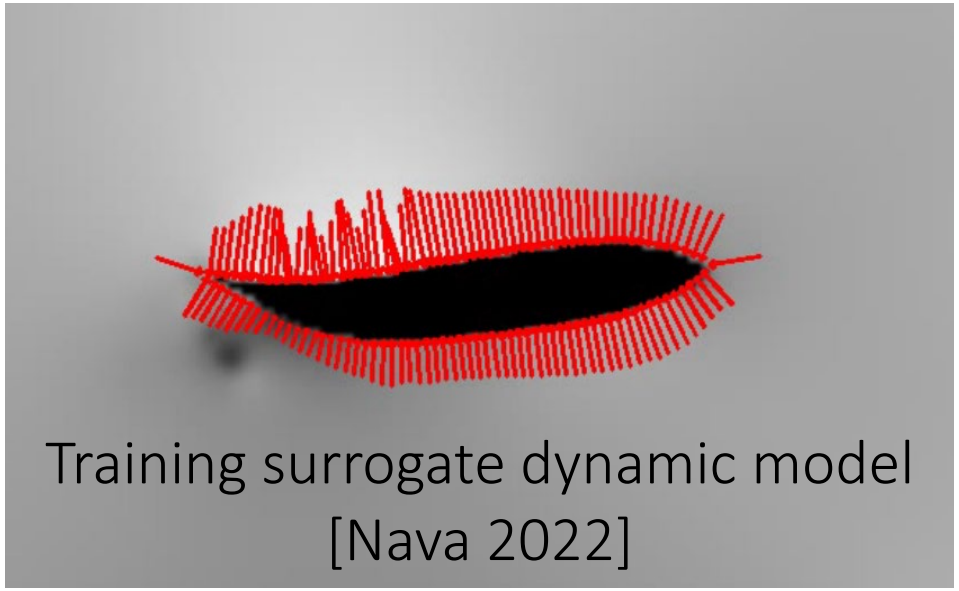
Goal: estimating scene parameters to reconstruct the balls' motion.



User gallery: machine learning

DiffPD also attracts users from the learning community.





Ma et al. *RISP: Rendering-Invariant State Predictor with Differentiable Simulation and Rendering for Cross-Domain Parameter Estimation*. ICLR 2022
Nava et al. *Fast Aquatic Swimmer Optimization with Differentiable Projective Dynamics and Neural Network Hydrodynamic Models*. ICML 2022

Summary

Conclusions

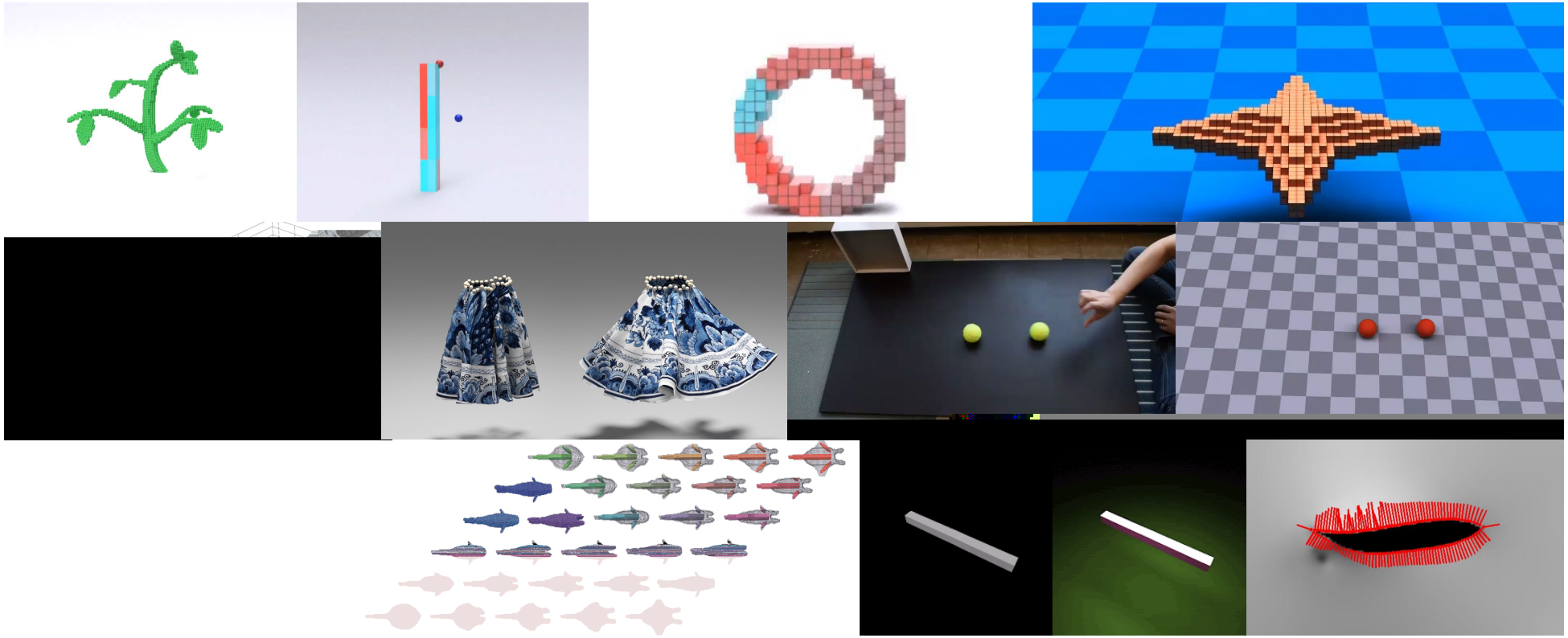
Numerical techniques in forward simulation and backpropagation are two sides of the same coin.

Efficient forward simulation: $\mathcal{P}\left(\mathcal{O}^{\mathcal{S}}\right)\mathcal{F}\mathcal{O}^{\mathcal{S}} = \mathcal{P}\left(\mathcal{O}^{\mathcal{S}}\right).$

Efficient backpropagation: $\mathcal{P}(\mathcal{O}_{\mathcal{S}\mathcal{T}\mathcal{P}})\mathcal{O} = \left(\frac{\mathcal{J}}{\mathcal{O}_{\mathcal{I}}}\right)^{\mathcal{T}}.$

Conclusions

A fast, reliable differentiable soft-body simulator unlocks wide application in graphics, robotics, and machine learning.



For more information

Project

http://diffpd.csail.mit.edu/

Code

https://github.com/mit-gfx/diff_pd_public

This work is sponsored by DARPA FA8750-20-C-0075, IARPA 2019-19020100001, and NSF 2106962.