Avoiding Poor Generalizability of Differentiable Simulation

Bibit Bianchini, Mathew Halm, Michael Posa

Penn Engineering | GRASP Laboratory
General Robotics, Automation, Sensing & Perception Lab

ROBOTICS SCIENCE AND SYSTEMS

PROBLEM FORMULATION

For model building, we aim to learn a function $y = f(x)$ with structure such that it can be formulated as an implicit optimization problem:

$$y = g(x, \lambda)$$
$$\text{such that} \quad \lambda = \underset{\lambda \in \Lambda}{\arg\min}\, h(x, y, \lambda)$$

This class can represent smooth and discontinuous functions alike. Specifically, this structure naturally accommodates rigid body contact dynamics where λ represents frictional forces maximizing energy dissipation.

We compare three approaches defined by the loss functions:

- **End-to-End Learning:** explicitly learn f with prediction loss.

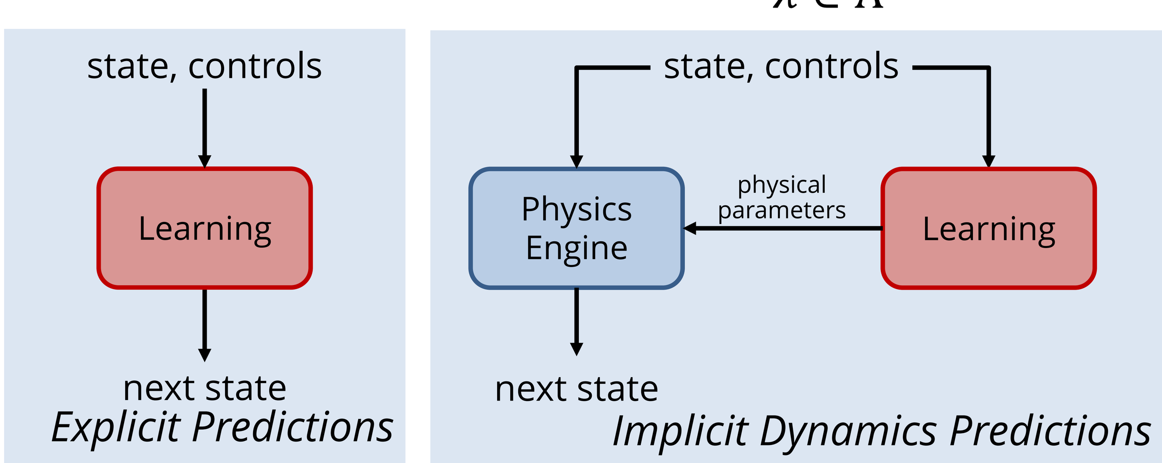
$$l^{\theta}_{\text{exp}}(x_i, y_i) = \left\| y_i - f^{\theta}(x_i) \right\|^2$$

- **Differentiable Simulation:** implicitly learn g with prediction loss. This is how differentiable simulation is formulated.

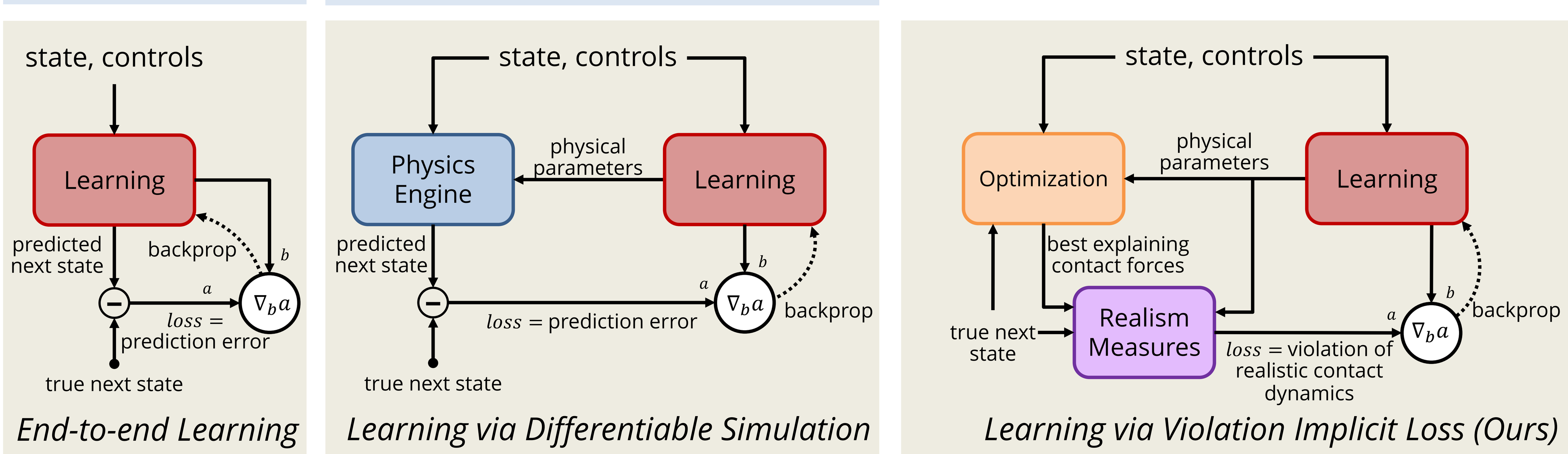
$$l^{\theta}_{\text{diffsim}}(x_i, y_i) = \left\| y_i - g^{\theta}(x_i, \lambda_i) \right\|^2$$
$$\text{such that} \quad \lambda_i = \underset{\lambda \in \Lambda}{\arg\min}\, h^{\theta}(x_i, g^{\theta}(x_i, \lambda), \lambda)$$

- ***Ours:* Violation Implicit Loss:** implicitly learn g with a loss that trades off between prediction loss and violation of constraints, requiring no simulation during training.

$$l^{\theta}_{\text{vimp}}(x_i, y_i) = \min_{\lambda \in \Lambda} \left\| y_i - g^{\theta}(x_i, \lambda) \right\|^2 + \frac{1}{\epsilon} h^{\theta}(x_i, y_i, \lambda)$$



We can prove that if the learned parameterization θ is shared, **end-to-end learning is equivalent to differentiable simulation** [Bianchini et al, L4DC 2022].



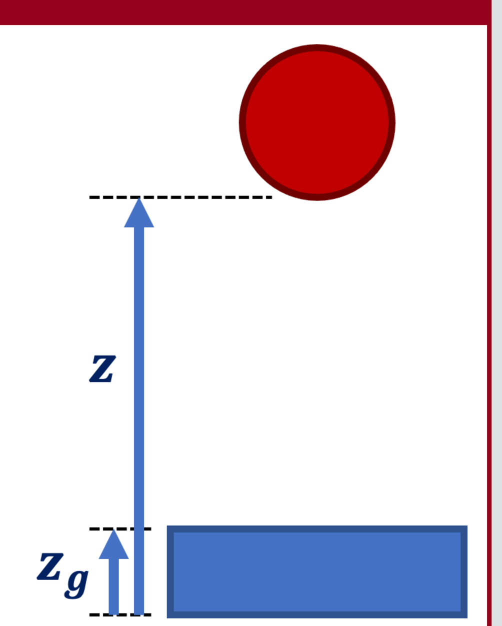
GENERALIZATION ERROR BOUNDS, TOY PROBLEM EXAMPLE

We ground our analysis with a physically-relevant 1-D toy problem of a point mass inelastically colliding with a rigid ground.

All approaches share the parameterization of learning the scalar value of the ground height, z_g.

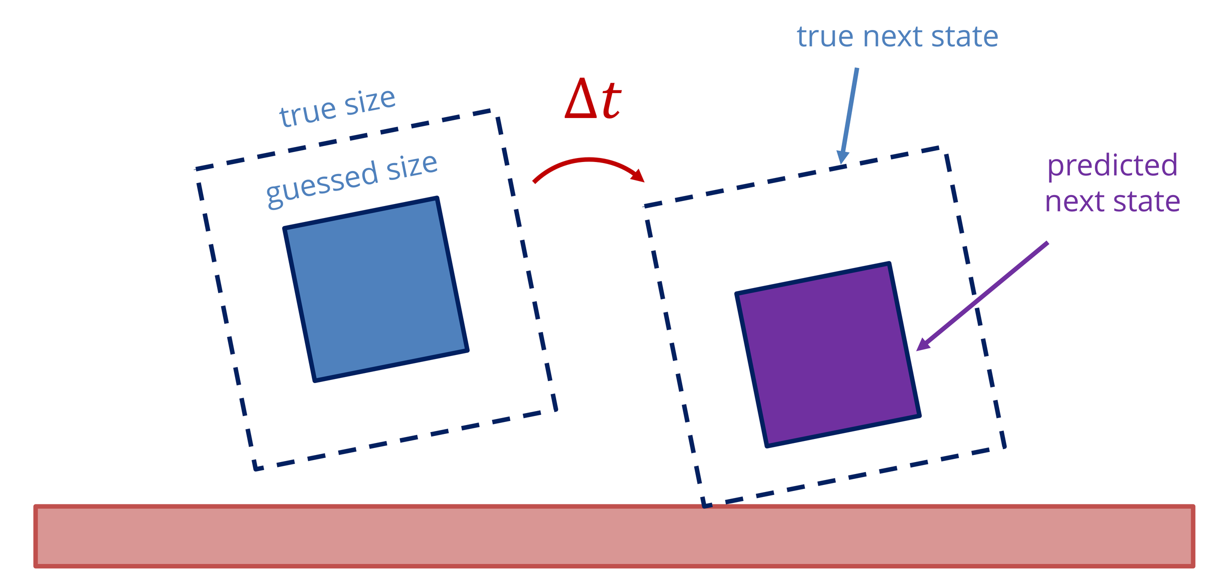
We combine prior works using Rademacher complexity and Dudley's entropy integral to bound generalization error for our approaches.

On our toy problem, **the generalization error bound scales poorly like $1/\Delta t$ for Differentiable Simulation, and independently of Δt for our Violation Implicit.**

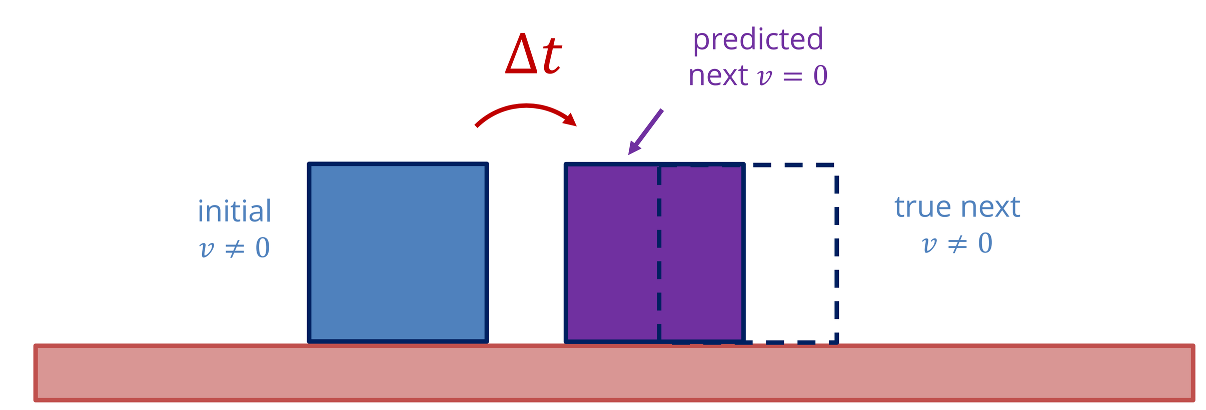


DIFFICULT SCENARIOS FOR DIFFERENTIABLE SIMULATION

If guessed geometry is too small such that contact is not predicted during single timesteps, the loss gradient w.r.t. contact parameters (e.g. geometry, friction) is 0 and the true parameters are unrecoverable.



If guessed friction is too high, the loss gradient w.r.t. the friction coefficient approaches zero with worse and worse overestimates.



Existing differentiable simulators can get around these flat loss regions by softening the contact model. Our Violation Implicit approach **achieves a smoother loss landscape while still modeling stiff contact dynamics exactly**.

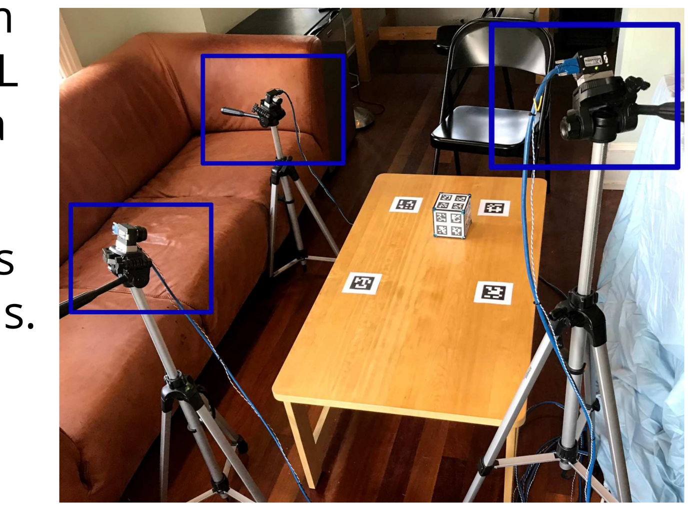
FUTURE WORK

Future work will explore:

- the practical implications of contact dynamics softening.
- incorporating system tracking without relying on AprilTags.
- how to further increase data efficiency by leveraging vision.

CONNECT WITH BIBIT

EMPIRICAL RESULTS ON CUBE TOSS DATA

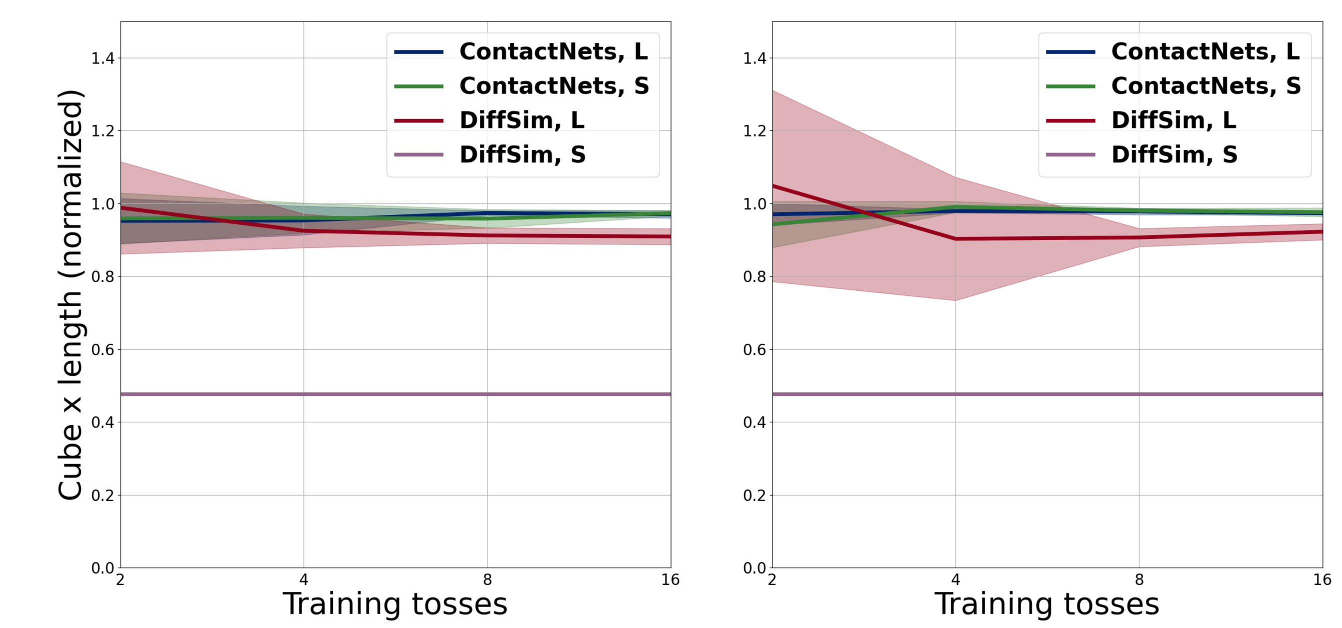


We use the cube toss dataset from ContactNets [Pfrommer et al, CoRL 2020] to learn the dimensions of a cube.

We train on single-step predictions and cut off training after 10 epochs.

We try 2 initialization scenarios:

- Geometry too large (L)
- Geometry too small (S)



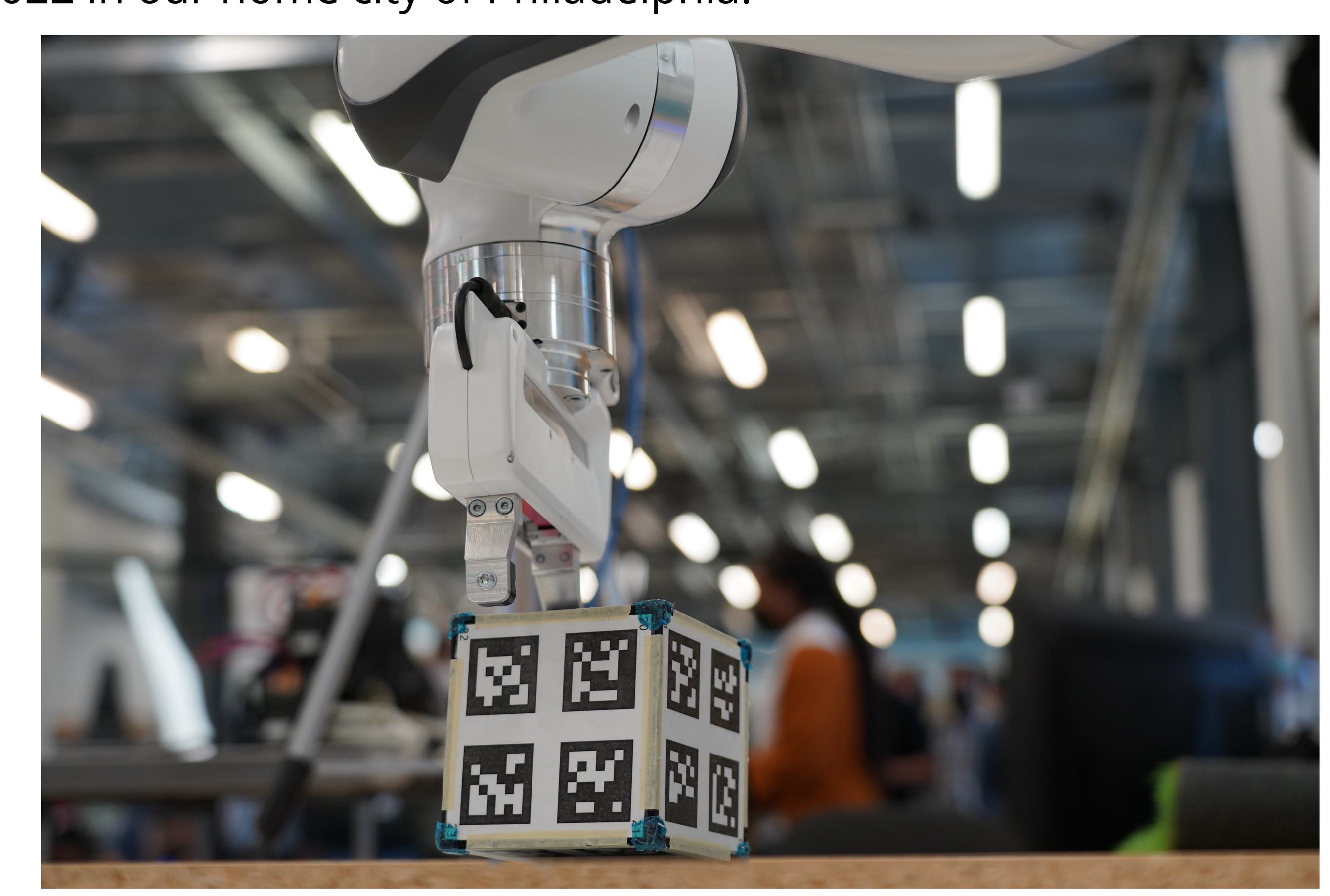
When geometry is initialized to be too small, **Differentiable Simulation cannot recover the true parameters.**

ContactNets trained with a Violation Implicit loss efficiently and accurately obtains the correct values, no matter the initial guess.

HIGH ENOUGH DATA EFFICIENCY FOR LIVE MODEL-BUILDING DEMO

End-to-end, Differentiable Simulation, and ContactNets approaches alike will work well with a big enough dataset. Our interest lies in low data regimes such that a robot could make sense of a newly encountered object in its goal to complete a task.

ContactNets is data efficient enough to run in real-time, exemplified by our live cube toss and model building demo on display at ICRA 2022 in our home city of Philadelphia.



bibit@seas.upenn.edu

This work was supported by an NDSEG Fellowship. Toyota Research Institute also provided funds for this work.