

---

# Taccel: Scaling Up Vision-based Tactile Robotics via High-performance GPU Simulation

---

Yuyang Li<sup>1,2,3,4,\*</sup>

y.li@stu.pku.edu.cn

Wenxin Du<sup>5,\*</sup>

setsunainn@gmail.com

Chang Yu<sup>5,\*</sup>

changyu1@g.ucla.edu

Puhao Li<sup>4</sup>

liph23@mails.tsinghua.edu.cn

Zihang Zhao<sup>1,2,3</sup>

zhaozihang@stu.pku.edu.cn

Tengyu Liu<sup>4</sup>

liutengyu@bigai.ai

Chenfanfu Jiang<sup>5,✉</sup>

cffjiang@math.ucla.edu

Yixin Zhu<sup>1,2,3,6,✉</sup>

yixin.zhu@pku.edu.cn

Siyuan Huang<sup>4,✉</sup>

huangsiyuan@ucla.edu

\*Equal contributor ✉Corresponding authors

<sup>1</sup> Institute for Artificial Intelligence, Peking University

<sup>2</sup> School of Psychological and Cognitive Sciences, Peking University

<sup>3</sup> Beijing Key Laboratory of Behavior and Mental Health, Peking University

<sup>4</sup> State Key Lab of General AI, Beijing Institute for General AI

<sup>5</sup> AIVC Laboratory, University of California, Los Angeles

<sup>6</sup> Embodied Intelligence Lab, PKU-Wuhan Institute for Artificial Intelligence

<https://taccel-simulator.github.io>

## Abstract

Tactile sensing is crucial for achieving human-level robotic capabilities in manipulation tasks [54]. As a promising solution, Vision-based tactile sensors (VBTs) [61, 37] offer high spatial resolution and cost-effectiveness, but present unique challenges in robotics for their complex physical characteristics and visual signal processing requirements. The lack of efficient and accurate simulation tools for VBTs has significantly limited the scale and scope of tactile robotics research [51, 8]. We present **Taccel**, a high-performance simulation platform that integrates Incremental Potential Contact (IPC) and Affine Body Dynamics (ABD) to model robots, tactile sensors, and objects with both accuracy and unprecedented speed, achieving a total of 915 FPS with 4096 parallel environments. Unlike previous simulators that operate at sub-real-time speeds with limited parallelization, **Taccel** provides precise physics simulation and realistic tactile signals while supporting flexible robot-sensor configurations through user-friendly APIs. Through extensive validation in object recognition, robotic grasping, and articulated object manipulation, we demonstrate precise simulation and successful sim-to-real transfer. These capabilities position **Taccel** as a powerful tool for scaling up tactile robotics research and development, potentially transforming how robots interact with and understand their physical environment.

## 1 Introduction

The ability to physically interact with the environment through touch is fundamental to robotic manipulation [3, 13]. While vision provides global scene understanding, tactile sensing captures crucial local contact information [57] essential for precise manipulation. Among various tactile sensing technologies [64, 25, 38, 26], vision-based tactile sensors (VBTs) such as GelSight [61] and 9DTact [37] have emerged as a central focus in tactile research. Their ability to provide high-resolution tactile feedback through camera-captured deformation patterns of elastic gel pads, combined with cost-effectiveness, has driven significant advances in robotics [66, 35, 47, 8, 65].

The primary challenge in scaling up VBT-equipped robot simulation lies in accurately modeling the hyperelastic soft gel pad and its contact [66, 12]. Current approaches follow two main directions: 39th Conference on Neural Information Processing Systems (NeurIPS 2025).

rigid-body approximations [51, 56] and soft-body simulations [47, 8, 12, 21, 30, 66]. While rigid-body methods efficiently support basic tasks like pick-and-place [1, 51], they cannot capture the fine-grained contact and elastomer deformations essential for complex manipulation tasks [66, 12] and detailed force distribution analysis [40, 47]. Soft-body simulations offer higher fidelity but face significant computational challenges that limit their practical application in large-scale experiments.

An ideal VBTS simulator must simultaneously achieve:

- **Precision:** Precise modeling of robots, sensors, and objects with physically valid solutions, particularly maintaining inversion-free and intersection-free states during complex contact interactions; generation of realistic tactile signals across multiple resolutions, from high-resolution RGB patterns and depth maps to low-resolution marker movements.
- **Scalability:** Capability for large-scale parallelization for extensive simulation data generation.
- **Flexibility:** Support for diverse robotic platforms and sensor configurations, from parallel grippers to multi-finger hands with varying sensor arrangements.

As detailed in Tab. 1, existing solutions often compromise on precision, scalability, or flexibility. They typically produce suboptimal physics, operate slower than real-time with limited parallel environments, or focus on specific sensor setups or simple grippers. These limitations significantly impede the broader application of tactile robotics.

To address these challenges, we present **Tacel**, a high-performance simulation platform for scaling up robots with VBTS-integration. Built on state-of-the-art simulation techniques, **Tacel** provides dedicated components for simulating robots (Sec. 4), tactile sensors (Sec. 4.2), and tactile signal generation (Sec. 4.3). Comprehensive evaluations (Sec. 5) demonstrate its characteristics:

- **Precision:** **Tacel** leverages advanced solid material simulation techniques (IPC [32] and ABD [31]) for physical accuracy. IPC guarantees inversion- and intersection-free contact solutions, while the integration of ABD allows for efficient and precise simulation.
- **Scalability:** With an efficient ABD-IPC implementation with NVIDIA Warp [41], **Tacel** achieves unprecedented parallelization. On a single H100 GPU, it reaches over 900 FPS in total (4096 environments,  $18\times$  wallclock time) for a peg-insertion task with dual sensors and 12.67 FPS (256 environments,  $0.25\times$  wallclock time) in a dexterous grasping task with full-hand tactile sensing.
- **Flexibility:** **Tacel** provides user-friendly APIs for seamless integration of diverse robotic platforms and sensor configurations. Users can easily load and configure robots through Unified Robot Description Format (URDF) with auxiliary configurations, supporting applications from simple grippers to complex manipulation tasks, like the mahjong tile sensing task in Fig. 1.

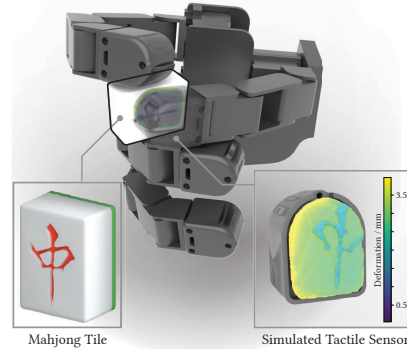


Figure 1: **Tacel demonstration of tactile robotics simulation.** An Allegro Hand with four VBTSs performing a precision grasp on a mahjong tile. The deformation map precisely captures the tile’s surface geometries.

Table 1: **Comprehensive comparison of FEM-based VBTS simulators.** **Soft Mat.:** modeling of deformable materials (FEM: Finite Element Methods). **Stiff Mat.:** modeling of stiff materials (Rigid: rigid body, ABD: Affine Body Dynamics, MPM: Material Point Methods, PBD: Position-based Dynamics). **Contact:** collision handling method (Virtual: approximated contact, Penalty: penalty-based, IPC: Incremental Potential Contact). **RGB Signal:** RGB tactile pattern generation method (Look-up: look-up tables, DNN: Deep Neural Network,  $\times$ : not supported). **Robot:** range of supported robotic systems. The last two columns report parallel simulation capabilities: maximum parallel environments and simulation speed relative to real-time in a peg insertion test, with dual sensors in low/high resolutions, measured on an NVIDIA H100 80G GPU. Details are in Fig. 4.

Simulator	Soft Mat.	Stiff Mat.	Contact	RGB Signal	Robot	# Env $\uparrow$	Sim Speed $\uparrow$
Taxim [46]	-	Rigid	Virtual	Look-up	Sensor	1	-
DiffTactile [47]	FEM	MPM/PBD	Penalty	DNN	Gripper	1	-
SAPIEN-IPC [8]	FEM	ABD	IPC	$\times$	Gripper	256 / 4	$0.81\times / 0.03\times$
<b>Tacel (Ours)</b>	FEM	ABD	IPC	DNN	Any	<b>4096 / 64</b>	<b><math>18.30\times / 0.25\times</math></b>

**Taccel**'s utility is validated through three fundamental tactile-informed robotic tasks (Sec. 6). In object classification, models trained solely on **Taccel**'s synthetic tactile signals demonstrate strong generalization to real-world data without adaptation. In grasping experiments across four robotic hand designs, we showcase the platform's versatility in handling diverse robot configurations and tactile signal types. In articulated object manipulation tasks, we demonstrate **Taccel**'s physical fidelity through close correspondence between simulated and real-world robot behavior.

Our key contributions include: (i) development of a high-performance simulation platform combining precise physics, realistic tactile signal generation, and massive parallelization; (ii) user-friendly APIs enabling flexible robot-sensor integration and high-fidelity tactile signal synthesis; (iii) comprehensive evaluation of the platform's precision and scalability; and (iv) extensive experimental validation across diverse tactile robotic tasks. By enabling large-scale, high-fidelity simulation of VBTS-equipped robots, **Taccel** aims to accelerate future research in tactile robotics.

## 2 Related Work

Early VBTS simulators focused on normal deformation scenarios, approximating hyperelastic behavior through geometric computations and surface modifications [17, 51, 56, 1, 46]. While efficient in generating high-resolution tactile signals through physics-based rendering or look-up tables, these approaches inadequately capture elastomer dynamics during complex manipulation tasks, especially those involving tangential forces and continuous interactions [66].

Recent approaches have achieved higher physical fidelity by incorporating advanced solid material simulation techniques. Methods using Material Point Methods (MPM) [49, 22, 23] and Finite Element Methods (FEM) [32, 31] better model elastomer properties through time-integrated deformation computations [10, 8, 47, 12]. Notable improvements include the adoption of IPC [32] by several simulators [12, 8], providing robust contact handling with guaranteed inversion- and intersection-free solutions. Tab. 1 compares key features of representative approaches.

**Taccel** builds on these advances by combining IPC and ABD in a unified platform, achieving both physical accuracy and computational efficiency while supporting diverse robot configurations and enabling large-scale parallel simulation for robot learning applications.

A more comprehensive review on related works is in Appx. B concerning the VBTSs and tactile-informed robotic tasks.

## 3 Unified IPC Simulation in Taccel

This section presents the unified IPC simulation framework in **Taccel**, detailing its mathematical foundations. For complete derivations, we refer readers to Sec. 3 and the original works [31, 32, 9].

### 3.1 Problem Formulation and Soft Body Dynamics

We consider  $n_s$  tetrahedralized soft bodies discretized into  $N_s$  vertices with positions  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{N_s}$  in Cartesian space. The system state is represented by the stacked position vector  $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_{N_s}^T]^T \in \mathbb{R}^{3N_s}$ . Following Lagrangian mechanics, we express the system's Lagrangian as  $L(\mathbf{x}, \dot{\mathbf{x}}) = T(\mathbf{x}, \dot{\mathbf{x}}) - V(\mathbf{x})$ , where  $T(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2} \dot{\mathbf{x}}^T \mathbf{M} \dot{\mathbf{x}}$  represents kinetic energy with mass matrix  $\mathbf{M} \in \mathbb{R}^{3N_s \times 3N_s}$ . The potential energy  $V(\mathbf{x})$  comprises two terms: an elastic energy  $\Phi(\mathbf{x})$  utilizing the Neo-Hookean constitutive model for hyperelastic materials (characterized by Young's modulus  $\mathcal{E}$  and Poisson's ratio  $\nu$ ), and external forces  $E_{\text{ext}}(\mathbf{x})$ .

### 3.2 Frictional Contact

The Euler - Lagrange equation  $\frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}, \dot{\mathbf{x}}) - \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}) = 0$  is equivalent to the following Incremental Potential (IP) energy minimization problem [32] under the backward Euler integration scheme:

$$\begin{aligned} E_{\text{IP}}(\mathbf{x}) &= \frac{1}{2} (\mathbf{x} - \mathbf{x}^n - \Delta t \dot{\mathbf{x}}^n)^T \mathbf{M} (\mathbf{x} - \mathbf{x}^n - \Delta t \dot{\mathbf{x}}^n) + \Delta t^2 V(\mathbf{x}), \\ \mathbf{x}^{n+1} &= \arg \min_{\mathbf{x}} E_{\text{IP}}(\mathbf{x}), \end{aligned} \tag{1}$$

where time is discretized into steps  $\{t_n = n\Delta t : n \in \mathbb{N}\}$  with a fixed step size  $\Delta t > 0$ , and  $\mathbf{x}^n = \mathbf{x}(t_n)$ .

To ensure intersection-free trajectories, IPC augments the objective with log barrier functions  $b(d_k(\mathbf{x}))$ , which diverge to  $+\infty$  as the distance  $d_k(\mathbf{x})$  between contact primitive pair  $k$  approaches zero. Additionally, IPC introduces an approximate frictional potential energy  $D_k(\mathbf{x}, \mathbf{x}^n)$  that captures frictional forces through its gradient. The full simulation thus minimizes the IPC energy:

$$E_{\text{IPC}}(\mathbf{x}) = E_{\text{IP}}(\mathbf{x}) + \Delta t^2 B(\mathbf{x}) + \Delta t^2 D(\mathbf{x}, \mathbf{x}^n). \quad (2)$$

with  $B(\mathbf{x}) = \kappa \sum_{k \in \mathcal{B}} A_k b(d_k(\mathbf{x}))$ ,  $D(\mathbf{x}, \mathbf{x}^n) = \sum_{k \in \mathcal{B}} D_k(\mathbf{x}, \mathbf{x}^n)$ , where  $\kappa > 0$  controls contact stiffness.

### 3.3 ABD and Unified Simulation

For  $n_a$  affine bodies, we introduce a reduced coordinate space  $\mathbf{y} \in \mathbb{R}^{12n_a}$  with an embedding map  $\phi : \mathbb{R}^{12n_a} \rightarrow \mathbb{R}^{3N_a}$  that projects reduced coordinates to full-space vertices  $\phi(\mathbf{y})$  [31], where  $N_a$  denotes the total vertex count of affine bodies' surface meshes. Each affine body uses 12 Degree of Freedom (DoF): three for translation ( $\mathbb{R}^3$ ) and nine for affine deformation ( $\mathbb{R}^{3 \times 3}$ ).

$$T(\mathbf{y}, \dot{\mathbf{y}}) = \frac{1}{2} \dot{\mathbf{x}}^T \mathbf{M} \dot{\mathbf{x}} = \frac{1}{2} \dot{\phi}(\mathbf{y})^T \mathbf{M} \dot{\phi}(\mathbf{y}) = \frac{1}{2} (\mathbf{J} \dot{\mathbf{y}})^T \mathbf{M} (\mathbf{J} \dot{\mathbf{y}}) = \frac{1}{2} \dot{\mathbf{y}}^T (\mathbf{J}^T \mathbf{M} \mathbf{J}) \dot{\mathbf{y}} = \frac{1}{2} \dot{\mathbf{y}}^T \mathbf{M}^y \dot{\mathbf{y}}, \quad (3)$$

where  $\mathbf{J} = \frac{\partial \phi}{\partial \mathbf{y}} \in \mathbb{R}^{3N_a \times 12n_a}$  is the Jacobian,  $\mathbf{M}$  is the full-space mass matrix, and  $\mathbf{M}^y = \mathbf{J}^T \mathbf{M} \mathbf{J}$  is the reduced-space mass matrix. The potential energy  $V(\mathbf{y})$  includes an As-Rigid-As-Possible (ARAP) term  $\Phi^y(\mathbf{y}) = \Phi^x(\phi(\mathbf{x}))$  with large  $\kappa_s$  to limit deformation, plus external forces  $E_{\text{ext}}(\mathbf{y})$ .

Combining with Eq. (2) yields the unified IPC energy [9] for the full system state  $\{\mathbf{y}; \mathbf{x}\} \in \mathbb{R}^{12n_a + 3N_s}$ .

$$\begin{aligned} E_{\text{IPC}}(\mathbf{y}; \mathbf{x}) &= E_{\text{IP}}(\mathbf{x}) + E_{\text{IP}}(\mathbf{y}) + \Delta t^2 B(\phi(\mathbf{y}); \mathbf{x}) + \Delta t^2 D(\phi(\mathbf{y}); \mathbf{x}, \phi(\mathbf{y}^n); \mathbf{x}^n), \\ E_{\text{IP}}(\mathbf{y}) &= \frac{1}{2} (\mathbf{y} - \mathbf{y}^n - \Delta t \dot{\mathbf{y}}^n)^T \mathbf{M}^y (\mathbf{y} - \mathbf{y}^n - \Delta t \dot{\mathbf{y}}^n) + \Delta t^2 V(\mathbf{y}). \end{aligned} \quad (4)$$

The next timestep's configuration follows from minimizing the following barrier-augmented IP:

$$\mathbf{y}^{n+1}; \mathbf{x}^{n+1} = \arg \min_{\mathbf{y}; \mathbf{x}} E_{\text{IPC}}(\mathbf{y}; \mathbf{x}). \quad (5)$$

### 3.4 Kinematic Constraints

We express kinematic constraints as  $\mathbf{S}^x \mathbf{x} = \mathbf{s}^x$  and  $\mathbf{S}^y \mathbf{y} = \mathbf{s}^y$ , where  $\mathbf{S}^x \in \mathbb{R}^{c^x \times 3N_s}$ ,  $\mathbf{s}^x \in \mathbb{R}^{3N_s}$  for soft bodies, and  $\mathbf{S}^y \in \mathbb{R}^{c^y \times 12n_a}$ ,  $\mathbf{s}^y \in \mathbb{R}^{12n_a}$  for affine bodies. To enforce the constraints, we augment  $E_{\text{IPC}}$  using the Augmented Lagrangian method with Lagrangian multipliers  $\lambda^x \in \mathbb{R}^{c^x}$ ,  $\lambda^y \in \mathbb{R}^{c^y}$ :

$$E_{\text{IPC}}^{\text{AL}}(\mathbf{y}; \mathbf{x}) = E_{\text{IPC}}(\mathbf{y}; \mathbf{x}) + \|(\mathbf{S}^x \mathbf{x} - \mathbf{s}^x)^T \lambda^x\|_2^2 + \|(\mathbf{S}^y \mathbf{y} - \mathbf{s}^y)^T \lambda^y\|_2^2. \quad (6)$$

## 4 Robot and VBTS Simulation in Taccel

Building upon the unified ABD-IPC, **Taccel** implements robot and VBTS simulation through a modular design that leverages the complementary strengths of affine- and soft-body dynamics.

### 4.1 Robot and Sensor Simulation

**Robot Modeling** For a robot with  $D$ -DoF,  $L$  links, and  $N$  vision-based tactile sensors (VBTSs), **Taccel** constructs its kinematic model from a URDF specification, loading visual and collision meshes as affine bodies (Sec. 3.3). Given the robot's global transformation  $T_r$  and joint configuration  $q \in \mathcal{Q}$ , forward kinematics yields the transformation of each link  $j$  in the world frame,  ${}^l_j T(q)$ .

**Tactile Sensors Modeling** Each VBTS is simulated by a tetrahedral mesh as a soft volumetric body (Sec. 3) for its gel pad, attached to its corresponding robot link. For the  $i$ -th sensor's gel pad  $\mathcal{G}_i$  attached to link  $l_j$  with local transformation  ${}^{l_j}_{\mathcal{G}_i} T$ , we denote its outer surface as  $\mathcal{S}_i = \partial \mathcal{G}$ . The surface comprises a reflective-coated region  $\partial^+ \mathcal{G}_i$  and a sensor-attached region  $\partial^- \mathcal{G}$ . Contact interactions during simulation cause gel pad deformation, transforming the coated surface  $\partial^+ \mathcal{G}$  to  $\partial^+ \tilde{\mathcal{G}}$  and marker positions to  $\tilde{\mathbf{P}}_i$ . These deformed quantities serve as the foundation for generating multiple types of tactile signals, each suited for different robotic applications.

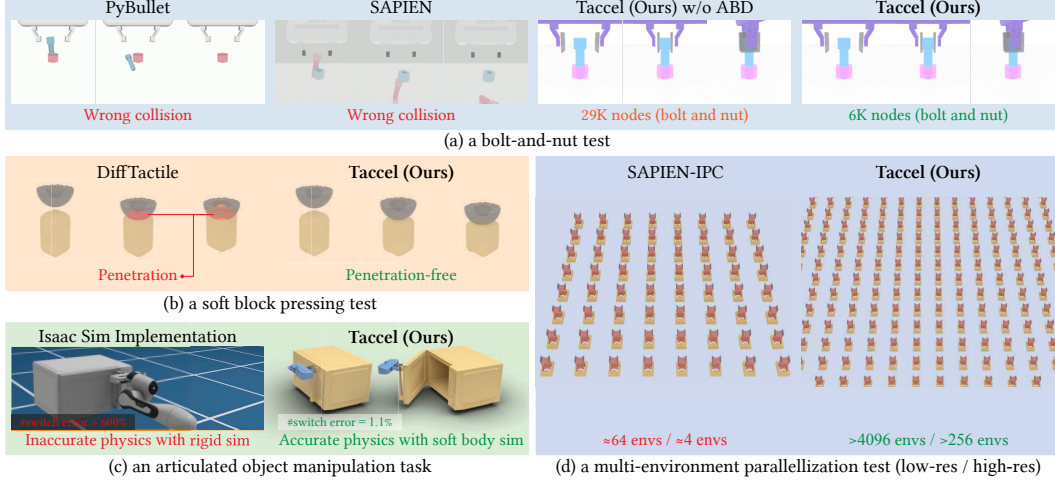


Figure 2: **Comprehensive evaluation of physics simulation capabilities across VBTS simulators.** (a) Bolt-nut assembly involving contact between non-convex objects, where **Taccel** achieves stable simulation; (b) Soft block pressing test with soft-soft contacts, where **Taccel** maintains penetration-free interactions; (c) Tactile-informed articulated object manipulation, where **Taccel** replicates real-world interactions with  $\sim 1\%$  physical error; (d) Parallel environment test using a peg insertion task, showing **Taccel**'s scalability.

**Robot Actions** For scene initialization, we compute affine states through forward kinematics for robot links and explicit state specification for stiff objects. Gel pad node positions are transformed to the world frame, with all states written to  $\mathbf{x}$ ,  $\mathbf{y}$ , and velocities  $\dot{\mathbf{x}}$ ,  $\dot{\mathbf{y}}$  initialized to zero. Robot actions are implemented through kinematic constraints. From joint space targets, we compute affine state targets for links and node position targets for gel pad attached surfaces  $\partial^-\mathcal{G}$ , assembled as  $\mathbf{s}^x, \mathbf{s}^y$ . Selection matrices  $\mathbf{S}^x, \mathbf{S}^y$  apply these constraints, with remaining states solved via time stepping.

## 4.2 Tactile Signal Simulation in Taccel

While many works directly generate tactile signals from the object geometry and frictional forces during rigid body simulation for efficiency [51, 56, 1], this approach can lead to inauthentic signals in dynamic scenarios. Instead, **Taccel** supports accurate simulation of high-resolution soft bodies to fully capture the fine-grained contact and deformation patterns.

**RGB Images and Depth Maps** High-resolution tactile signals, including RGB images and depth maps, are essential when fine-grained details like object texture and local geometries are required. The signal generation process proceeds in two stages. First, we extract the depth and normal maps  $d_{(u,v)}, n_{(u,v)}$  for pixel coordinates  $(u, v)$  from the deformed coated surface  $\partial^+\tilde{\mathcal{G}}$ . Next, following the method of Si *et al.* [47], we apply a Deep Neural Network (DNN) to generate RGB tactile signals from the depth information. Specifically, for each pixel coordinate  $(u, v)$ , a pixel-to-pixel DNN parameterized by  $\theta$  maps the inputs to the pixel colors relative to a reference image (RGB signal of undeformation gel pad):  $f_\theta(\gamma(u, v), n_{(u,v)}) \mapsto \Delta\sigma_{(u,v)}$ , which are added to the reference image to obtain the final RGB image.  $\gamma(\cdot, \cdot)$  provides the 2D positional encoding of the pixel coordinate. The model is trained on patches from 200 real tactile images and corresponding depth map annotations.

**Markers** Within  $\partial^+\mathcal{G}_i$ ,  $m_i$  markers are positioned at locations  ${}^{\mathcal{G}_i}\mathbf{P}_i = \{p_k^{(i)} \in \mathbb{R}^3, k = 1, \dots, m_i\}_{i=1}^N$ , each defined by barycentric coordinates in its triangle:  $p_k^{(i)} = \sum_{u=1}^3 \alpha_u \mathbf{x}_u^{(i,k)}$ , where  $\sum_{u=1}^3 \alpha_u = 1, \alpha_u \in [0, 1]$ . Low-resolution tactile signals primarily track marker positions and their movements by computing their new positions throughout the simulation (at time step  $t$ ):  $\tilde{p}_k^{(i)}(t) = \sum_{u=1}^3 \alpha_u \tilde{\mathbf{x}}_u^{(i,k)}(t)$ , and projecting them on the tactile image. The marker flows, representing local deformation patterns, are then computed as:  $\Delta\mathbf{P}_i = \{\tilde{p}_k^{(i)} - p_k^{(i)}\} = \{\Delta p_k^{(i)}\}$ .

**3D Tactile Signals** The depth map and marker positions can be transformed into a dense or sparse 3D point cloud in the world frame using robot kinematics and sensor configurations. These 3D tactile signals provide crucial spatial information for robotic manipulation tasks. We demonstrate their effectiveness through the simulation of Tac-Man framework [66] in Sec. 6.3.



We further demonstrate **Taccel**’s capability to scale up synthetic data generation of tactile signals via robotic grasping simulations and explore object recognition model learning (Secs. 6.1 and 6.2).

### 4.3 API Designs in Taccel

**Taccel** provides intuitive Python APIs designed to make tactile robotics simulation accessible to researchers while maintaining high performance through NVIDIA Warp [41]. The APIs allows for seamless loading of robots from URDF files with automatic parsing, sensor configurations from auxiliary files, and objects from mesh files. Users can efficiently reset simulation states or apply kinematic targets to control the robots in familiar formats (NumPy arrays, PyTorch tensors). Further, **Taccel** supports parallel simulation of multiple environments similar to Isaac Gym [43].

To foster community development, we will release **Taccel** codebase and documentations, while maintaining active collaboration with researchers to incorporate feedback, add features, and expand capabilities, ensuring its evolution as a comprehensive tool for tactile robotics research.

## 5 Performance Evaluation of Taccel

We evaluate **Taccel** through comprehensive benchmarks on its precision and efficiency. Our analysis demonstrates that the combination of ABD and IPC provides significant advantages over existing approaches (Sec. 5.1), generates high-quality tactile signals (Sec. 5.2), ensures precise frictions and deformation solving for the gel pad (Sec. 5.3), and enables efficient scaling (Sec. 5.4).

### 5.1 Overall Comparisons

As illustrated in Fig. 2, **Taccel** achieves superior precision and efficiency through its unified ABD and IPC framework, demonstrated across four challenging scenarios.

First, the bolt screwing task (Fig. 2a) demonstrates **Taccel**’s ability to handle complex physical interactions between highly non-convex objects, where conventional simulators including PyBullet and SAPIEN [55] fail to handle. This capability stems from our ABD formulation, which provides an efficient approach to simulating dense rigid-soft body interactions. Alternative approaches either model stiff objects as soft bodies, introducing excessive DoFs and computational overhead (Fig. 2a, ours w/o ABD), or rely on RBD [14], which requires expensive nonlinear Continuous Collision Detection (CCD) calculations to avoid intersection, significantly degrading performance.

Next, physical realism in **Taccel** is demonstrated through the soft block pressing scenario (Fig. 2b), where our collision-free and intersection-free guarantees produce notably more realistic deformations compared to penalty-based approaches. This precision extends to practical robotics applications, as shown in the Tac-Man microwave manipulation task (Fig. 2c). Here, **Taccel**’s accurate contact force solving enables faithful reproduction of gel pad-object handle interactions, closely matching real-world execution patterns (detailed analysis in Sec. 6.3).

Finally, the computational efficiency of **Taccel** emerges from our optimized implementation of the ABD and IPC algorithms, enabling unprecedented scaling capabilities. In a peg insertion task, **Taccel** achieves parallel simulation of over 4096 environments on a single GPU with 80GB VRAM—representing a 64-fold improvement over SAPIEN-IPC [8]. This opens new possibilities for large-scale robotics simulation and learning; see also the comprehensive analysis in Sec. 5.4.

### 5.2 Tactile Signal Simulation

We evaluate the fidelity of tactile signals generated by **Taccel** with real-world samples. We press a calibrated GelSight-type sensor perpendicularly on 18 objects from a standard tactile shape testing dataset [16] on a real-world setup (Fig. 3(a)) and in **Taccel**; see Appx. D.1 for setup details.

The qualitative and quantitative comparisons shown in Fig. 3(a) demonstrate **Taccel**’s ability to produce highly realistic tactile patterns, achieving an average SSIM of 0.93 across all test objects. Minor variations between simulated and real signals primarily stem from manufacturing tolerances in the 3D-printed objects and challenges in precise camera calibration. Despite these practical limitations, the results establish **Taccel**’s capability to generate high-fidelity tactile signals suitable for VBTSSs, with simulated patterns closely matching experimental measurements.

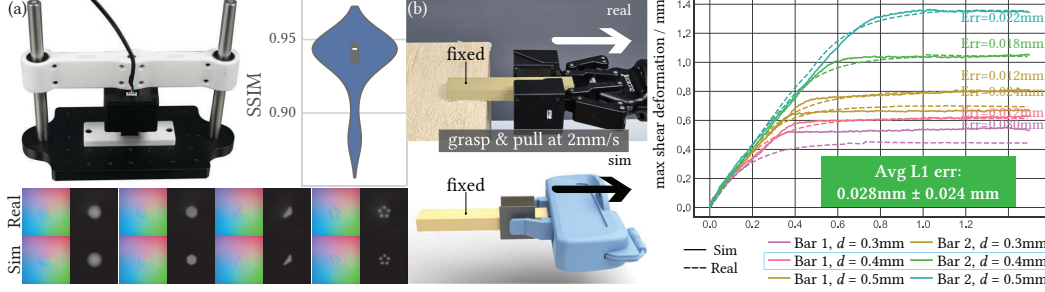


Figure 3: **Evaluations on the simulation precision.** (a) Data collection setup, examples of the real and simulated tactile patterns, and the sim-real SSIM distribution (violin plot) in the tactile signal evaluation. (b) Data collection setup and the shear deformation magnitude trajectories in the frictions simulation evaluation.

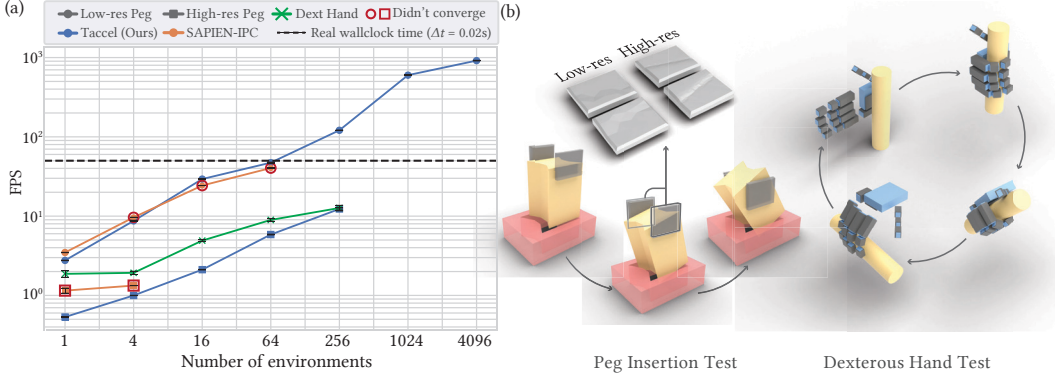


Figure 4: **Parallel simulation performance analysis across environment scaling.** (a) FPS achieved by **Taccel** (FP64) and SAPIEN-IPC (FP32) on an NVIDIA H100 80G GPU. (b) Task visualizations. [Appx. C.1](#) reports the VRAM occupancy in the low-resolution peg-insertion test.

### 5.3 Precision on Frictions and Shear Deformation

We further investigate **Taccel**’s precision on solving frictions and shear deformation. [Fig. 3\(b\)](#) shows the evaluation setup: A VBTS-gripper grasps a fixed bar with various forces (causes tactile depth  $d$ ) then pulls back at 2 mm/s, recording gel deformation represented by the marker deformations. We calibrate the object’s friction coefficient with one record and use two others with various  $d$ s to compare the errors between simulated and real-world trajectories. Small sim-real errors (avg 28  $\mu$ m) tested on two bars with various frictions highlights **Taccel**’s fidelity on gel deformation and frictions in both static and slipping cases.

### 5.4 Multi-environment Simulation

Efficient and stable parallel simulation of multiple environments is crucial for scaling up synthetic data collection across diverse downstream tasks. To evaluate **Taccel**’s capabilities in this regard, we designed three test cases of increasing complexity: (i) a dual-sensor (139 nodes per gel pad) peg-insertion task adapted from SAPIEN-IPC [8], (ii) the peg-insertion task with higher solution (1.5k nodes per gel pad), and (iii) a grasping task using a customized five-fingered dexterous hand equipped with 17 gel pads wrapped around the hand links (5k nodes). These tasks involve continuous contact and gel pad deformation, shedding light on how the simulator would perform on various robotic tasks. [Appx. D.2](#) provide more details on these tasks.

With a single NVIDIA H100 80G GPU, we benchmarked **Taccel** against SAPIEN-IPC. The results in [Fig. 4](#) showcase **Taccel**’s superior performance: in the low-resolution peg-insertion task, **Taccel** achieves 915 FPS (18.30 $\times$  wallclock time) while managing over 4096 parallel environments—a 16-fold improvement over the baseline using the same GPU memory, enabled by our efficient parallelization implementation. The high-resolution test further demonstrates **Taccel**’s advantages, maintaining both stability and precision while consistently outperforming the baseline. Even in the complex dexterous hand scenario, **Taccel** efficiently handles full-hand tactile sensing, achieving 12.67 FPS across 256 environments.

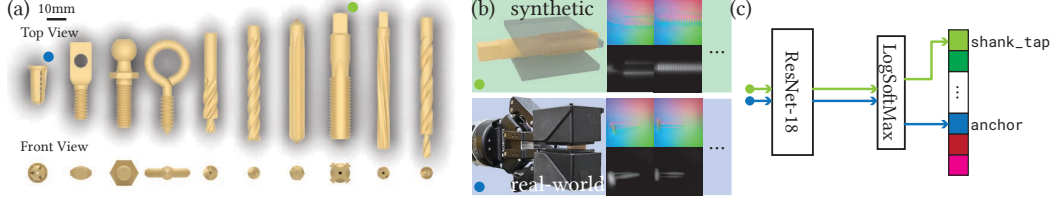


Figure 5: **Object classification experiment setups.** (a) 10 mechanical parts selected, each with diverse and fine-grained geometries, rendered from the top and the front. (b) Synthetic and real-world data collection setup. (c) A small ResNet is learned to perform object classification.

We observed that SAPIEN-IPC’s use of FP32 precision leads to convergence issues when solving contact forces in Eq. (2), particularly in the logarithmic barrier energy term calculations, as indicated by red outlines in Fig. 4. We further investigate this gap in the low-resolution peg-insertion benchmark. We set the maximum Newton iterations to 50 and optimization residue tolerance set to 0.01m/s, *i.e.* the IPC solver performs conjugate gradient descent step until (i) the 50-step limit is exceeded, or (ii) the optimization residue is below 0.01m/s. We report the key quantities including the Newton iterations used, its optimization residue, total time spent for the simulation, and total FPS in Tab. 2. Although each step takes longer for FP64, the convergence is much faster due to the better precision, and thus the better simulation speed.

Noteworthy, while **Tacel**’s FP64 precision is ideal for HPC GPUs (which have a high 1:2 FP32:FP64 FLOPS ratio), it remains highly performant on more accessible consumer cards like RTX 3090 and 4090; see Appx. C.2 for more details.

Table 2: **Simulation Performance Comparison.** Results are grouped by the number of parallel environments. The number of Newton iterations (# Newton Its.) and optimization residue (Optim. Residue) are reported as mean with standard deviation (in parentheses), calculated across the 200-step simulation.

Simulator	# Envs	# Newton Its. ↓	Optim. Residue / (m/s) ↓	Time / s ↓	Total FPS ↑
<b>Tacel</b> (FP64)	1	<b>1.59 (±1.13)</b>	$7.64 \times 10^{-5} (\pm 6.55 \times 10^{-5})$	<b>69.78</b>	<b>2.87</b>
SAPIEN-IPC (FP32)	1	24.09 (±10.35)	0.001 (±0.01)	82.41	2.43
<b>Tacel</b> (FP64)	16	<b>2.83 (± 2.57)</b>	$1.87 \times 10^{-4} (\pm 2.09 \times 10^{-4})$	<b>141.92</b>	<b>22.55</b>
SAPIEN-IPC (FP32)	16	50.0 (±0.0)	0.11 (±0.44)	291.46	10.98
<b>Tacel</b> (FP64)	64	<b>3.45 (±3.26)</b>	$2.04 \times 10^{-4} (\pm 2.20 \times 10^{-4})$	<b>196.71</b>	<b>65.07</b>
SAPIEN-IPC (FP32)	64	50.0 (±0.0)	0.20 (±0.44)	668.95	19.13

## 6 Applications in Tactile Robotics

We demonstrate **Tacel**’s capabilities across three tactile-informed robotic tasks: training object classification models with synthetic data (Sec. 6.1), generating a large-scale dataset through parallel grasp simulations (Sec. 6.2), and manipulating articulated objects (Sec. 6.3). These applications showcase how **Tacel** enables precise robotic simulation and scalable synthetic data generation.

### 6.1 Learning Object Classification Models

To demonstrate **Tacel**’s ability to generate synthetic training data that generalizes to real-world scenarios, we developed a tactile-based object classification system. As shown in Fig. 5, our approach trains a DNN to classify objects using high-resolution tactile signals.

Table 3: **Object classification and sim-to-real performance.** Performance metrics include number of tactile sensors ( $N$ ), Degree of Freedom (DoF), total sensing area (S.A.), average sensor-object contact area (C.A.) with percentage of total sensing area in parentheses, and classification accuracy (Acc.).

(a) Sim-to-real results on mechanical part recognition.						(b) Results on various robotic hand configurations.					
Data	DoF	$N$	S.A. / cm <sup>2</sup>	C.A. / cm <sup>2</sup>	Acc. ↑	Data	DoF	$N$	S.A. / cm <sup>2</sup>	C.A. / cm <sup>2</sup>	Acc. ↑
Mech.	2	2	32.0	4.77 (14.92%)	86.50%	Gripper	2	2	32.0	5.05 (15.78%)	44.56%
Mech. (Real)	2	2	32.0	5.69 (17.78%)	70.94%	Robotiq-3F	8	3	27.0	4.98 (18.43%)	44.61%
						Allegro	16	4	23.0	<b>7.67 (33.34%)</b>	<b>54.30%</b>
						F-TAC [65]	15	17	59.7	4.00 (6.700%)	42.54%



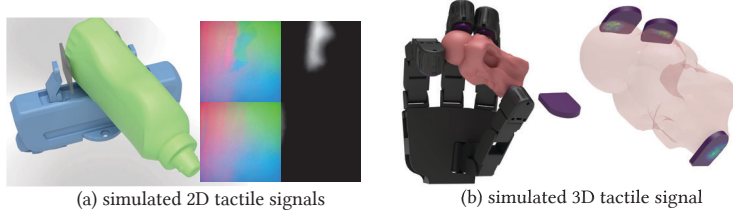


Figure 6: **Examples of the synthesized grasps and the simulated tactile signals.**

Following Yang *et al.* [58], we selected 10 mechanical parts with distinct fine-grained geometries (illustrated in Fig. 5, top). We collect a training dataset 4K tactile depth maps of a tactile sensor pressing on them within **Taccel**, each with object pose and tactile depth randomly sampled. We also collect a real-world test set that consists of 160 depth maps for testing. Appx. D.3 illustrates more details for the data collection protocol. We trained a ResNet-18 model [20] for 10-category object classification using the simulated depth images. To enhance robustness, we augmented the depth maps through random affine transformations, morphological operations (erosion and dilation), and Gaussian filtering. The model is then directly evaluated them on the real-world samples, with each sample tested 4 times with random shear deformation. As shown in Tab. 3, our model achieved 86.50% accuracy on the synthetic test set and 70.94% on real-world samples without any domain adaptation. This modest sim-to-real gap demonstrates **Taccel**’s capability to generate precise tactile signals that enable data-efficient training of transferable tactile perception models.

## 6.2 Robotic Grasping with Tactile Sensors

We investigate robotic grasping across different hand configurations with varying tactile sensor arrangements. We first extend the DFC algorithm [39] for generating contact-oriented grasping poses for four robotic hands, then simulate their tactile responses within **Taccel**, as illustrated in Fig. 6. We generate grasps on 10 diverse objects from ContactDB [4], YCB [7], and adversarial object [42] datasets. For each object, we generated grasps using 4 different robotic hands, producing  $\sim 14k$  total grasps. These were simulated in **Taccel** to generate tactile signals, with key metrics summarized in Tab. 3. Fig. A1 provides an additional visualizations for the grasps across 4 robotic hands. Appx. D.4 explains the details of the modification to DFC.

Simulating robotic grasps with tactile sensors yields synthetic data extending the existing robotic grasping datasets with tactile perception capabilities, serving as a foundation for various robotic tasks. To demonstrate its utility, we implemented the object classification task described in Sec. 6.1, adapting it to use the deformed coat’s point cloud as input and PointNet as the feature extractor.

The classification performances in Tab. 3 reveals an interesting trade-off: while robots differ in sensor count and sensing area, higher dexterity (Allegro Hand) enables better object contact despite fewer sensors. This enhanced contact leads to superior classification accuracy, highlighting the balance between sensor count and dexterity in tactile hand design. These findings demonstrate **Taccel**’s value in validating robotic hand designs before physical fabrication.

## 6.3 Articulated Object Manipulation

We consider articulated object manipulation, a challenging task where tactile perception provides critical feedback on hand-object contact, informing object articulation and guiding robot actions [26, 66]. Appx. D.5 detailedly explain the algorithm.

Tac-Man’s effectiveness relies heavily on gel pad deformation for motion adaptation and tactile feedback. While the original implementation by Zhao *et al.* used rigid body simulation with gripper compliance approximations, it couldn’t authentically replicate gel pad deformation, contact dynamics, and tactile feedback, resulting in significant sim-to-real gaps during large-scale verification. **Taccel** overcomes these limitations through accurate simulation of sensor-object contact and gel pad deformation. We simulate Tac-Man on three types of articulated objects: drawers with prismatic joints, cabinets with revolute joints, and bolt-nut pairs with helical joints. Fig. 7(a) shows the manipulation sequences and corresponding tactile signals.

To evaluate sim-real correspondence, we compared our simulation against real-world Tac-Man execution using a microwave (revolute joint) and drawer (prismatic joint), shown in Fig. 7(b-c). For

comprehensive comparison, we also tested Tac-Man’s official implementation on these objects. We use the “execution-recovery switch count” as our key metric, which tallies how often the agent switch between the execution and recovery state when executing the Tac-Man algorithm (see Zhao *et al.* [66] and Appx. D.5 for details). As demonstrated in Fig. 7, **Taccel** faithfully reproduces real-world execution patterns, with execution-recovery switch counts averaging 68.75 and 0.0 for revolute and prismatic settings, respectively—remarkably matching real-world observations. While Isaac Sim successfully simulated the manipulation, its dynamics gap resulted in substantially higher switch counts. These results highlight **Taccel**’s capability to authentically replicate physical interactions in manipulation scenarios.

## 7 Summary

We present **Taccel**, a flexible and high-performance simulator for VBTS-integrated robots. Its user-friendly APIs enables precise and efficient simulation of complex tactile robotic tasks with realistic tactile signals. **Taccel** excels in capturing intricate deformation and contact dynamics of soft gel pads with unprecedented stability, while supporting thousands of parallelized environments. These capabilities also position **Taccel** as a powerful tool for hand-sensor validating before fabrication, potentially reducing development time and costs.

**Limitations** Despite the high performance of **Taccel**, its computational demands of large-scale simulation remain challenging, with a major bottleneck being the PCG-based linear system solving. Possible remedies include carefully relaxing convergence tolerances, simplifying simulation protocols, and using larger timesteps, enabled by IPCs unconditionally stable solver. Besides, the Neo-Hookean model itself may still be insufficient in capturing all necessary physical characteristics of the gel pad. Further, the data-driven tactile signal generation method is still trained from limited data and thus may not generalize to all contact conditions. These limitations are to be addressed in future studies.

**Broader Impact** This paper focuses on fundamental research, with limited direct societal impact.

## Acknowledgments and Disclosure of Funding

This work is supported in part by the National Science and Technology Innovation 2030 Major Program (2025ZD0219402), the National Natural Science Foundation of China (62376009), the Beijing Nova Program, the State Key Lab of General AI at Peking University, the PKU-BingJi Joint Laboratory for Artificial Intelligence, and the National Comprehensive Experimental Base for Governance of Intelligent Society, Wuhan East Lake High-Tech Development Zone. We thank Lei Yan (LeapZenith AI Research) for his support in mechanical engineering, Zhenghao Qi (PKU, THU), Sirui Xie (PKU) for their assistance in setting up experiments, and Ms. Hailu Yang (PKU) for her help in procuring raw materials.

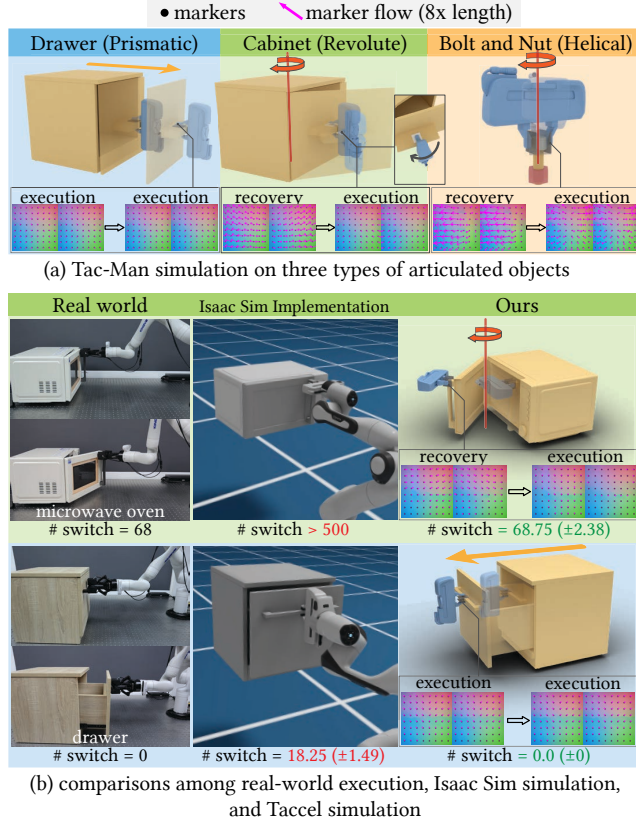


Figure 7: **Tac-Man manipulation simulation.** (a) Demonstration of Tac-Man’s execution-recovery cycles on three articulated objects. (b) Three-way comparison among real-world execution, Isaac Sim implementation from Zhao *et al.* [66], and **Taccel**. The execution-recovery switch counts (#switch) demonstrate **Taccel**’s accuracy in replicating real-world behavior (1.10% error), compared to Isaac Sim’s higher counts.

## References

- [1] Iretiayo Akinola, Jie Xu, Jan Carius, Dieter Fox, and Yashraj Narang. Tacs1: A library for visuotactile sensor simulation and learning. *IEEE Transactions on Robotics*, 2025. 2, 3, 5
- [2] Maria Bauza, Antonia Bronars, and Alberto Rodriguez. Tac2pose: Tactile object pose estimation from the first touch. *International Journal of Robotics Research (IJRR)*, 42(13):1185–1209, 2023. A3
- [3] Aude Billard and Danica Kragic. Trends and challenges in robot manipulation. *Science*, 364(6446):eaat8414, 2019. 1, A3
- [4] Samarth Brahmbhatt, Cusuh Ham, Charles C Kemp, and James Hays. Contactdb: Analyzing and predicting grasp contact via thermal imaging. In *Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019. 9
- [5] Antonia Bronars, Sangwoon Kim, Parag Patre, and Alberto Rodriguez. Texterity: Tactile extrinsic dexterity. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2024. A3
- [6] Roberto Calandra, Andrew Owens, Dinesh Jayaraman, Justin Lin, Wenzhen Yuan, Jitendra Malik, Edward H Adelson, and Sergey Levine. More than a feeling: Learning to grasp and regrasp using vision and touch. *IEEE Robotics and Automation Letters (RA-L)*, 3(4):3300–3307, 2018. A3
- [7] Berk Calli, Arjun Singh, Aaron Walsman, Siddhartha Srinivasa, Pieter Abbeel, and Aaron M Dollar. The ycb object and model set: Towards common benchmarks for manipulation research. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2015. 9
- [8] Weihang Chen, Jing Xu, Fanbo Xiang, Xiaodi Yuan, Hao Su, and Rui Chen. General-purpose sim2real protocol for learning contact-rich manipulation with marker-based visuotactile sensors. *IEEE Transactions on Robotics (T-RO)*, 40:1509–1526, 2024. 1, 2, 3, 6, 7, A3, A4
- [9] Yunuo Chen, Minchen Li, Lei Lan, Hao Su, Yin Yang, and Chenfanfu Jiang. A unified newton barrier method for multibody dynamics. *ACM Transactions on Graphics (TOG)*, 41(4):1–14, 2022. 3, 4, A2
- [10] Zixi Chen, Shixin Zhang, Shan Luo, Fuchun Sun, and Bin Fang. Tacchi: A pluggable and low computational cost elastomer deformation simulator for optical tactile sensors. *IEEE Robotics and Automation Letters (RA-L)*, 8(3):1239–1246, 2023. 3
- [11] Siyuan Dong, Wenzhen Yuan, and Edward H Adelson. Improved gelsight tactile sensor for measuring geometry and slip. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017. A3
- [12] Wenxin Du, Wenqiang Xu, Jieji Ren, Zhenjun Yu, and Cewu Lu. Tacipc: Intersection-and inversion-free fem-based elastomer simulation for optical tactile sensors. *IEEE Robotics and Automation Letters (RA-L)*, 2024. 1, 2, 3
- [13] Mark Edmonds, Feng Gao, Hangxin Liu, Xu Xie, Siyuan Qi, Brandon Rothrock, Yixin Zhu, Ying Nian Wu, Hongjing Lu, and Song-Chun Zhu. A tale of two explanations: Enhancing human trust by explaining robot behavior. *Science Robotics*, 4(37):eaay4663, 2019. 1
- [14] Zachary Ferguson, Minchen Li, Teseo Schneider, Francisca Gil-Ureta, Timothy Langlois, Chenfanfu Jiang, Denis Zorin, Danny M Kaufman, and Daniele Panozzo. Intersection-free rigid body dynamics. *ACM Transactions on Graphics (TOG)*, 40(4), 2021. 6
- [15] Carlo Ferrari, John F Canny, et al. Planning optimal grasps. In *IEEE International Conference on Robotics and Automation (ICRA)*, 1992. A3
- [16] Daniel Fernandes Gomes, Paolo Paoletti, and Shan Luo. Generation of gelsight tactile images for sim2real learning. *IEEE Robotics and Automation Letters (RA-L)*, 6(2):4177–4184, 2021. 6, A4

- [17] Daniel Fernandes Gomes, Achu Wilson, and Shan Luo. Gelsight simulation for sim2real learning. In *ICRA ViTac Workshop*, 2019. [3](#)
- [18] Phillip L Gould and Yuan Feng. *Introduction to linear elasticity*, volume 2. Springer, 1994. [A3](#)
- [19] Xiaofeng Guo, Hung-Jui Huang, and Wenzhen Yuan. Estimating properties of solid particles inside container using touch sensing. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023. [A3](#)
- [20] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016. [9](#)
- [21] Francois R Hogan, Jose Ballester, Siyuan Dong, and Alberto Rodriguez. Tactile dexterity: Manipulation primitives with tactile feedback. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2020. [2](#), [A3](#)
- [22] 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. [3](#)
- [23] Yuanming Hu, Jiancheng Liu, Andrew Spielberg, Joshua B Tenenbaum, William T Freeman, Jiajun Wu, Daniela Rus, and Wojciech Matusik. Chainqueen: A real-time differentiable physical simulator for soft robotics. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2019. [3](#)
- [24] Hung-Jui Huang, Xiaofeng Guo, and Wenzhen Yuan. Understanding dynamic tactile sensing for liquid property estimation. *Robotics: Science and Systems (RSS)*, 2022. [A3](#)
- [25] Bassam Jalgha, Daniel Asmar, Elie Shammas, and Imad Elhajj. Hierarchical fall avoidance strategy for small-scale humanoid robots. In *2012 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2012. [1](#)
- [26] Carlos A Jara, Jorge Pomares, Francisco A Candelas, and Fernando Torres. Control framework for dexterous manipulation using dynamic visual servoing and tactile sensors feedback. *Sensors*, 14(1):1787–1804, 2014. [1](#), [9](#)
- [27] Yousef Javanmardi, Huw Colin-York, Nicolas Szita, Marco Fritzsche, and Emad Moeendarbary. Quantifying cell-generated forces: Poissons ratio matters. *Communications physics*, 4(1):237, 2021. [A3](#)
- [28] Chunpeng Jiang, Wenqiang Xu, Yutong Li, Zhenjun Yu, Longchun Wang, Xiaotong Hu, Zhengyi Xie, Qingkun Liu, Bin Yang, Xiaolin Wang, et al. Capturing forceful interaction with deformable objects using a deep learning-powered stretchable tactile array. *Nature Communications*, 15(1):9513, 2024. [A3](#)
- [29] Roland S Johansson and J Randall Flanagan. Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nature Reviews Neuroscience*, 10(5):345–359, 2009. [A3](#)
- [30] Sangwoon Kim, Antonia Bronars, Parag Patre, and Alberto Rodriguez. Texterity–tactile extrinsic dexterity: Simultaneous tactile estimation and control for extrinsic dexterity. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2024. [2](#)
- [31] Lei Lan, Danny M Kaufman, Minchen Li, Chenfanfu Jiang, and Yin Yang. Affine body dynamics: fast, stable and intersection-free simulation of stiff materials. *ACM Transactions on Graphics (TOG)*, 41(4):1–14, 2022. [2](#), [3](#), [4](#), [A2](#)
- [32] Minchen Li, Zachary Ferguson, Teseo Schneider, Timothy R Langlois, Denis Zorin, Daniele Panozzo, Chenfanfu Jiang, and Danny M Kaufman. Incremental potential contact: intersection- and inversion-free, large-deformation dynamics. *ACM Transactions on Graphics (TOG)*, 39(4):49, 2020. [2](#), [3](#), [A1](#)

- [33] Mingxuan Li, Lunwei Zhang, Yen Hang Zhou, Tiemin Li, and Yao Jiang. Easycalib: Simple and low-cost in-situ calibration for force reconstruction with vision-based tactile sensors. *IEEE Robotics and Automation Letters*, 2024. **A3**
- [34] Rui Li, Robert Platt, Wenzhen Yuan, Andreas Ten Pas, Nathan Roscup, Mandayam A Srinivasan, and Edward Adelson. Localization and manipulation of small parts using gelsight tactile sensing. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2014. **A3**
- [35] Wanlin Li, Zihang Zhao, Leiyao Cui, Weiyi Zhang, Hangxin Liu, Li-An Li, and Yixin Zhu. Minitac: An ultra-compact 8 mm vision-based tactile sensor for enhanced palpation in robot-assisted minimally invasive surgery. *IEEE Robotics and Automation Letters (RA-L)*, 9(12):11170–11177, 2024. **1, A3**
- [36] Yuyang Li, Bo Liu, Yiran Geng, Puhao Li, Yaodong Yang, Yixin Zhu, Tengyu Liu, and Siyuan Huang. Grasp multiple objects with one hand. *IEEE Robotics and Automation Letters (RA-L)*, 9(5):4027–4034, 2024. **A3**
- [37] Changyi Lin, Han Zhang, Jikai Xu, Lei Wu, and Huazhe Xu. 9dtact: A compact vision-based tactile sensor for accurate 3d shape reconstruction and generalizable 6d force estimation. *IEEE Robotics and Automation Letters (RA-L)*, 2023. **1, A3**
- [38] Hongbin Liu, Juan Greco, Xiaojing Song, Joao Bimbo, Lakmal Seneviratne, and Kaspar Althoefer. Tactile image based contact shape recognition using neural network. In *2012 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 2012. **1**
- [39] Tengyu Liu, Zeyu Liu, Ziyuan Jiao, Yixin Zhu, and Song-Chun Zhu. Synthesizing diverse and physically stable grasps with arbitrary hand structures using differentiable force closure estimator. *IEEE Robotics and Automation Letters (RA-L)*, 7(1):470–477, 2021. **9, A3, A5**
- [40] John Lloyd and Nathan F Lepora. Pose-and-shear-based tactile servoing. *International Journal of Robotics Research (IJRR)*, 43(7):1024–1055, 2024. **2, A3**
- [41] Miles Macklin. Warp: A high-performance python framework for gpu simulation and graphics. <https://github.com/nvidia/warp>, March 2022. NVIDIA GPU Technology Conference (GTC). **2, 6**
- [42] Jeffrey Mahler, Jacky Liang, Sherdil Niyaz, Michael Laskey, Richard Doan, Xinyu Liu, Juan Aparicio Ojea, and Ken Goldberg. Dex-net 2.0: Deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics. In *Robotics: Science and Systems (RSS)*, 2017. **9**
- [43] Viktor Makoviychuk, Lukasz Wawrzyniak, Yunrong Guo, Michelle Lu, Kier Storey, Miles Macklin, David Hoeller, Nikita Rudin, Arthur Allshire, Ankur Handa, et al. Isaac gym: High performance gpu-based physics simulation for robot learning. *arXiv preprint arXiv:2108.10470*, 2021. **6**
- [44] Juntian Qu, Baijin Mao, Zhenkun Li, Yining Xu, Kunyu Zhou, Xiangyu Cao, Qigao Fan, Minyi Xu, Bin Liang, Houde Liu, et al. Recent progress in advanced tactile sensing technologies for soft grippers. *Advanced Functional Materials*, 33(41):2306249, 2023. **A3**
- [45] Yu She, Shaoxiong Wang, Siyuan Dong, Neha Sunil, Alberto Rodriguez, and Edward Adelson. Cable manipulation with a tactile-reactive gripper. *International Journal of Robotics Research (IJRR)*, 40(12-14):1385–1401, 2021. **A3**
- [46] Zilin Si and Wenzhen Yuan. Taxim: An example-based simulation model for gelsight tactile sensors. *IEEE Robotics and Automation Letters (RA-L)*, 7(2):2361–2368, 2022. **2, 3**
- [47] Zilin Si, Gu Zhang, Qingwei Ben, Branden Romero, Zhou Xian, Chao Liu, and Chuang Gan. Diff tactile: A physics-based differentiable tactile simulator for contact-rich robotic manipulation. In *Proceedings of International Conference on Learning Representations (ICLR)*, 2024. **1, 2, 3, 5**



- [48] Bruno Siciliano and Oussama Khatib. Robotics and the handbook. In *Springer Handbook of Robotics*, pages 1–6. Springer, 2016. [A3](#)
- [49] Deborah Sulsky, Zhen Chen, and Howard L Schreyer. A particle method for history-dependent materials. *Computer methods in applied mechanics and engineering*, 118(1-2):179–196, 1994. [3](#)
- [50] Sudharshan Suresh, Haozhi Qi, Tingfan Wu, Taosha Fan, Luis Pineda, Mike Lambeta, Jitendra Malik, Mrinal Kalakrishnan, Roberto Calandra, Michael Kaess, et al. Neuralfeels with neural fields: Visuotactile perception for in-hand manipulation. *Science Robotics*, 9(96):eadl0628, 2024. [A3](#)
- [51] Shaoxiong Wang, Mike Lambeta, Po-Wei Chou, and Roberto Calandra. Tacto: A fast, flexible, and open-source simulator for high-resolution vision-based tactile sensors. *IEEE Robotics and Automation Letters (RA-L)*, 7(2):3930–3937, 2022. [1](#), [2](#), [3](#), [5](#)
- [52] Zhi Wang, Yuchen Mo, Shengmiao Jin, and Wenzhen Yuan. Doorbot: Closed-loop task planning and manipulation for door opening in the wild with haptic feedback. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2025. [A3](#)
- [53] Benjamin Ward-Cherrier, Nicholas Pestell, Luke Cramphorn, Benjamin Winstone, Maria Elena Giannaccini, Jonathan Rossiter, and Nathan F Lepora. The tactip family: Soft optical tactile sensors with 3d-printed biomimetic morphologies. *Soft Robotics*, 5(2):216–227, 2018. [A3](#)
- [54] Goran Westling and Roland S Johansson. Factors influencing the force control during precision grip. *Experimental Brain Research*, 53:277–284, 1984. [1](#), [A3](#)
- [55] Fanbo Xiang, Yuzhe Qin, Kaichun Mo, Yikuan Xia, Hao Zhu, Fangchen Liu, Minghua Liu, Hanxiao Jiang, Yifu Yuan, He Wang, et al. Sapien: A simulated part-based interactive environment. In *Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. [6](#)
- [56] Jie Xu, Sangwoon Kim, Tao Chen, Alberto Rodriguez Garcia, Pulkit Agrawal, Wojciech Matusik, and Shinjiro Sueda. Efficient tactile simulation with differentiability for robotic manipulation. In *Conference on Robot Learning (CoRL)*, 2023. [2](#), [3](#), [5](#)
- [57] Wenqiang Xu, Zhenjun Yu, Han Xue, Ruolin Ye, Siqiong Yao, and Cewu Lu. Visual-tactile sensing for in-hand object reconstruction. In *Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. [1](#), [A3](#)
- [58] Sanghoon Yang, Won Dong Kim, Hyunkyu Park, Seojung Min, Hyonyoung Han, and Jung Kim. In-hand object classification and pose estimation with sim-to-real tactile transfer for robotic manipulation. *IEEE Robotics and Automation Letters (RA-L)*, 9(1):659–666, 2023. [9](#), [A3](#)
- [59] Kunpeng Yao and Aude Billard. Exploiting kinematic redundancy for robotic grasping of multiple objects. *IEEE Transactions on Robotics (T-RO)*, 39(3):1982–2002, 2023. [A3](#)
- [60] Zhenjun Yu, Wenqiang Xu, Pengfei Xie, Yutong Li, and Cewu Lu. Dynamic reconstruction of hand-object interaction with distributed force-aware contact representation. *arXiv preprint arXiv:2411.09572*, 2024. [A3](#)
- [61] Wenzhen Yuan, Siyuan Dong, and Edward H Adelson. Gelsight: High-resolution robot tactile sensors for estimating geometry and force. *Sensors*, 17(12):2762, 2017. [1](#), [A3](#)
- [62] Wenzhen Yuan et al. *Tactile measurement with a gelsight sensor*. PhD thesis, Massachusetts Institute of Technology, 2014. [A3](#)
- [63] Wenzhen Yuan, Rui Li, Mandayam A Srinivasan, and Edward H Adelson. Measurement of shear and slip with a gelsight tactile sensor. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 304–311. IEEE, 2015. [A3](#)
- [64] Kappasov Zhanat, Juan-Antonio Corrales, and Perdereau Veronique. Tactile sensing in dexterous robot hands-review. *Robotics & Autonomous Systems*, 74:195–220, 2015. [1](#)

- [65] Zihang Zhao, Wanlin Li, Yuyang Li, Tengyu Liu, Boren Li, Meng Wang, Kai Du, Hangxin Liu, Yixin Zhu, Qining Wang, et al. Embedding high-resolution touch across robotic hands enables adaptive human-like grasping. *Nature Machine Intelligence*, pages 1–12, 2025. [1](#), [8](#)
- [66] Zihang Zhao, Yuyang Li, Wanlin Li, Zhenghao Qi, Lecheng Ruan, Yixin Zhu, and Kaspar Althoefer. Tac-Man: Tactile-informed prior-free manipulation of articulated objects. *IEEE Transactions on Robotics (T-RO)*, 41:538–557, 2024. [1](#), [2](#), [3](#), [5](#), [9](#), [10](#), [A3](#), [A4](#), [A6](#)
- [67] Yixin Zhu, Tao Gao, Lifeng Fan, Siyuan Huang, Mark Edmonds, Hangxin Liu, Feng Gao, Chi Zhang, Siyuan Qi, Ying Nian Wu, et al. Dark, beyond deep: A paradigm shift to cognitive ai with humanlike common sense. *Engineering*, 6(3):310–345, 2020. [A3](#)

## NeurIPS Paper Checklist

### 1. Claims

Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

Answer: [\[Yes\]](#)

Justification: The abstract and introduction clearly state the claims made.

Guidelines:

- The answer NA means that the abstract and introduction do not include the claims made in the paper.
- The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
- The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
- It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

### 2. Limitations

Question: Does the paper discuss the limitations of the work performed by the authors?

Answer: [\[Yes\]](#)

Justification: Section [Sec. 7](#) includes a "Limitations" paragraph.

Guidelines:

- The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
- The authors are encouraged to create a separate "Limitations" section in their paper.
- The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

### 3. Theory assumptions and proofs

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [\[Yes\]](#)

Justification: The derivation of the unified IPC-ABD algorithm is included in [Sec. 3](#).

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

#### 4. Experimental result reproducibility

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [\[Yes\]](#)

Justification: The paper discloses all the required information for the simulation system and the experiments. The codebase is ready and will be released upon acceptance for better reproducibility.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general, releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  - (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  - (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  - (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  - (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

#### 5. Open access to data and code

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

Justification: The codebase and dataset are ready and will be released upon acceptance for better reproducibility.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so No is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

## 6. Experimental setting/details

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [Yes]

Justification: Most of the key details are provided within the text. Full details will be provided within the code.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

## 7. Experiment statistical significance

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification: The simulation speed test in Sec. 5 (see Fig. 4) includes the error bars. We also report the standard deviation for the Tac-Man simulation in Fig. 7.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.



- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

#### 8. Experiments compute resources

Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer: [Yes]

Justification: The paper includes necessary information on the computer resources.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

#### 9. Code of ethics

Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics <https://neurips.cc/public/EthicsGuidelines>?

Answer: [Yes]

Justification: The research conducted in the paper conforms with the NeurIPS Code of Ethics.

Guidelines:

- The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

#### 10. Broader impacts

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [Yes]

Justification: The broader impact is discussed in [Sec. 7](#).

Guidelines:

- The answer NA means that there is no societal impact of the work performed.

- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

#### 11. Safeguards

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: The paper does now contain data or models with a high risk for misuse.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

#### 12. Licenses for existing assets

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Justification: The authors of the assets are cited.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.

- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, [paperswithcode.com/datasets](https://paperswithcode.com/datasets) has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

### 13. New assets

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [NA]

Justification: The paper does not release new assets.

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

### 14. Crowdsourcing and research with human subjects

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [NA]

Justification: The paper does not contain crowdsourcing experiments or research with human subjects.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

### 15. Institutional review board (IRB) approvals or equivalent for research with human subjects

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Justification: The paper does not contain research with human subjects.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.

- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

#### 16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

Answer: [NA]

Justification: LLMs are only used for editing the manuscript. Thus the declaration is not needed.

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (<https://neurips.cc/Conferences/2025/LLM>) for what should or should not be described.

## A Unified ABD-IPC in Taccel

In **Taccel**, robot links are efficiently modeled as affine bodies to capture their primarily rigid motion, while VBTSSs are simulated as soft bodies to accurately represent their deformation mechanics. This natural division allows **Taccel** to balance computational efficiency with physical accuracy while maintaining consistent contact handling through IPC.

We provide a more detailed derivation of the unified ABD-IPC algorithm underlying **Taccel**, as previously briefed in [Sec. 3](#).

### A.1 Soft Body Dynamics

Recall that we consider  $n_s$  tetrahedralized soft bodies with  $N_s$  vertices:  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{N_s}$ . The stacked position vector  $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_{N_s}^T]^T \in \mathbb{R}^{3N_s}$  represents the system state, and thus the system's Lagrangian is  $L(\mathbf{x}, \dot{\mathbf{x}}) = T(\mathbf{x}, \dot{\mathbf{x}}) - V(\mathbf{x})$  with the kinetic energy  $T(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2} \dot{\mathbf{x}}^T \mathbf{M} \dot{\mathbf{x}}$ , given the mass matrix  $\mathbf{M} \in \mathbb{R}^{3N_s \times 3N_s}$ . The potential energy  $V(\mathbf{x})$  is composed of the elastic energy  $\Phi(\mathbf{x})$  utilizing the Neo-Hookean constitutive model for hyperelastic materials (characterized by Young's modulus  $\mathcal{E}$  and Poisson's ratio  $\nu$ ), and the external forces  $E_{\text{ext}}(\mathbf{x})$ . The elastic energy is defined as  $\Phi(\mathbf{x}) = \int_{\Omega} \Psi(\mathbf{x}) d\mathbf{x}$ , where  $\Psi(\mathbf{x})$  denotes elastic energy density over the volume region  $\Omega$  of all objects in rest configuration.

### A.2 Time Stepping

Substituting  $L(\mathbf{x}, \dot{\mathbf{x}})$  into the Euler-Lagrange equation  $\frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}, \dot{\mathbf{x}}) - \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}) = 0$  yields the governing dynamics:

$$\mathbf{M} \ddot{\mathbf{x}} = - \frac{dV}{d\mathbf{x}}(\mathbf{x}). \quad (\text{A1})$$

We temporally discretize [Eq. \(A1\)](#) using backward Euler:

$$\frac{\mathbf{x}^{n+1} - \mathbf{x}^n}{\Delta t} = \dot{\mathbf{x}}^{n+1}, \quad \frac{\mathbf{M}(\dot{\mathbf{x}}^{n+1} - \dot{\mathbf{x}}^n)}{\Delta t} = - \frac{dV}{d\mathbf{x}}(\mathbf{x}^{n+1}), \quad (\text{A2})$$

where time is discretized into steps  $\{t_n = n\Delta t : n \in \mathbb{N}\}$  with step size  $\Delta t > 0$ , and  $\mathbf{x}^n = \mathbf{x}(t_n)$ . Under this discretization, [Eq. \(A1\)](#) can be formulated as:

$$\frac{d}{d\mathbf{x}} (E_{\text{IP}}(\mathbf{x}^n)) = 0. \quad (\text{A3})$$

If we define the incremental potential energy of the constrained system as:

$$E_{\text{IP}}(\mathbf{x}) = \frac{1}{2} (\mathbf{x} - \mathbf{x}^n - \Delta t \dot{\mathbf{x}}^n)^T \mathbf{M} (\mathbf{x} - \mathbf{x}^n - \Delta t \dot{\mathbf{x}}^n) + \Delta t^2 V(\mathbf{x}), \quad (\text{A4})$$

then the general simulation problem in a conservative system can be reformulated as the minimization problem:

$$\mathbf{x}^{n+1} = \arg \min_{\mathbf{x}} E_{\text{IP}}(\mathbf{x}). \quad (\text{A5})$$

### A.3 Frictional Contact

We employ IPC [\[32\]](#) to handle contact interactions. The method operates on surface contact pairs  $\mathcal{B}$ , comprising point-triangle and edge-edge pairs from the surface meshes of soft and affine objects. For each contact pair  $k \in \mathcal{B}$  with distance  $d_k > 0$ , IPC defines two key energy terms. First, a barrier energy that prevents interpenetration:

$$b(d_k(\mathbf{x})) = - \left( d_k - \hat{d} \right)^2 \log \left( \frac{d_k}{\hat{d}} \right) I_{\{d_k \in (0, \hat{d})\}}(d_k), \quad (\text{A6})$$

where  $\hat{d} > 0$  is the distance threshold for contact force activation and  $I(\cdot)$  is the indicator function.



Second, an approximated friction potential energy:

$$D_k(\mathbf{x}, \mathbf{x}^n) = \mu \lambda_k^n f_0(\|\mathbf{u}_k\|), \quad (\text{A7})$$

where  $\mathbf{x}^n$  represents the configuration at the previous timestep  $t_n$ ,  $\lambda_k^n$  is the magnitude of the lagged normal contact force, and  $\mathbf{u}_k \in \mathbb{R}^2$  denotes the tangential relative displacement in the local contact frame.

The friction transition function  $f_0(x) = \int_{\epsilon_v \Delta t}^x f_1(y) dy + \epsilon_v \Delta t$  uses:

$$f_1(y) = \begin{cases} -\frac{y^2}{\epsilon_v^2 \Delta t^2} + \frac{2y}{\epsilon_v \Delta t}, & y \in (0, \Delta t \epsilon_v), \\ 1, & y \geq \Delta t \epsilon_v, \end{cases} \quad (\text{A8})$$

where  $\epsilon_v > 0$  serves as a velocity threshold distinguishing between static and dynamic friction regimes.

These contact and friction terms augment our incremental potential energy:

$$E_{\text{IPC}}(\mathbf{x}) = E_{\text{IP}}(\mathbf{x}) + \Delta t^2 B(\mathbf{x}) + \Delta t^2 D(\mathbf{x}, \mathbf{x}^n), \quad (\text{A9})$$

with  $B(\mathbf{x}) = \kappa \sum_{k \in \mathcal{B}} A_k b(d_k(\mathbf{x}))$ ,  $D(\mathbf{x}, \mathbf{x}^n) = \sum_{k \in \mathcal{B}} D_k(\mathbf{x}, \mathbf{x}^n)$ , where  $\kappa > 0$  controls contact stiffness.

#### A.4 ABD and Unified Simulation

For  $n_a$  affine bodies, we introduce a reduced coordinate space  $\mathbf{y} \in \mathbb{R}^{12n_a}$  with an embedding map  $\phi: \mathbb{R}^{12n_a} \rightarrow \mathbb{R}^{3N_a}$  that projects reduced coordinates to full-space vertices  $\phi(\mathbf{y})$  [31], where  $N_a$  denotes the total vertex count of affine bodies' surface meshes. Each affine body uses 12 DoF: three for translation ( $\mathbb{R}^3$ ) and nine for affine deformation ( $\mathbb{R}^{3 \times 3}$ ).

$$\begin{aligned} T(\mathbf{y}, \dot{\mathbf{y}}) &= \frac{1}{2} \dot{\mathbf{x}}^T \mathbf{M} \dot{\mathbf{x}} = \frac{1}{2} \dot{\phi}(\mathbf{y})^T \mathbf{M} \dot{\phi}(\mathbf{y}) \\ &= \frac{1}{2} (\mathbf{J} \dot{\mathbf{y}})^T \mathbf{M} (\mathbf{J} \dot{\mathbf{y}}) = \frac{1}{2} \dot{\mathbf{y}}^T (\mathbf{J}^T \mathbf{M} \mathbf{J}) \dot{\mathbf{y}} = \frac{1}{2} \dot{\mathbf{y}}^T \mathbf{M}^y \dot{\mathbf{y}}, \end{aligned} \quad (\text{A10})$$

where  $\mathbf{J} = \frac{\partial \phi}{\partial \mathbf{y}} \in \mathbb{R}^{3N_a \times 12n_a}$  is the Jacobian,  $\mathbf{M}$  is the full-space mass matrix, and  $\mathbf{M}^y = \mathbf{J}^T \mathbf{M} \mathbf{J}$  is the reduced-space mass matrix. The potential energy  $V(\mathbf{y})$  includes an As-Rigid-As-Possible (ARAP) term  $\Phi^y(\mathbf{y}) = \Phi^x(\phi(\mathbf{x}))$  with high stiffness  $\kappa_s$  to limit deformation, plus external forces  $E_{\text{ext}}(\mathbf{y})$ .

Combining with Eq. (A9), we obtain the unified affine-deformable coupled IPC energy [9] for the full system state  $\{\mathbf{y}; \mathbf{x}\} \in \mathbb{R}^{12n_a + 3N_s}$ :

$$\begin{aligned} E_{\text{IPC}}(\mathbf{y}; \mathbf{x}) &= E_{\text{IP}}(\mathbf{x}) + E_{\text{IP}}(\mathbf{y}) + \Delta t^2 B(\phi(\mathbf{y}); \mathbf{x}) \\ &\quad + \Delta t^2 D(\phi(\mathbf{y}); \mathbf{x}, \phi(\mathbf{y}^n); \mathbf{x}^n), \end{aligned} \quad (\text{A11})$$

where  $E_{\text{IP}}(\mathbf{y})$  is defined as:

$$\begin{aligned} E_{\text{IP}}(\mathbf{y}) &= \frac{1}{2} (\mathbf{y} - \mathbf{y}^n - \Delta t \dot{\mathbf{y}}^n)^T \mathbf{M}^y (\mathbf{y} - \mathbf{y}^n - \Delta t \dot{\mathbf{y}}^n) \\ &\quad + \Delta t^2 V(\mathbf{y}). \end{aligned} \quad (\text{A12})$$

The next timestep's configuration follows from minimizing this barrier-augmented incremental potential:

$$\mathbf{y}^{n+1}; \mathbf{x}^{n+1} = \arg \min_{\mathbf{y}; \mathbf{x}} E_{\text{IPC}}(\mathbf{y}; \mathbf{x}). \quad (\text{A13})$$

#### A.5 Kinematic Constraints

The kinematic constraints is expressed as

$$\mathbf{S}^x \mathbf{x} = \mathbf{s}^x, \mathbf{S}^y \mathbf{y} = \mathbf{s}^y, \quad (\text{A14})$$

where  $\mathbf{S}^x \in \mathbb{R}^{c^x \times 3N_s}$ ,  $\mathbf{s}^x \in \mathbb{R}^{3N_s}$  for soft bodies, and  $\mathbf{S}^y \in \mathbb{R}^{c^y \times 12n_a}$ ,  $\mathbf{s}^y \in \mathbb{R}^{12n_a}$  for affine bodies. The constraints are applied by selecting the constrained DoFs of the state vectors and specifying the

constraint values with. To enforce these constraints, we employ the Augmented Lagrangian method by augmenting  $E_{\text{IPC}}$  to:

$$E_{\text{IPC}}^{\text{AL}}(\mathbf{y}; \mathbf{x}) = E_{\text{IPC}}(\mathbf{y}; \mathbf{x}) + \|(\mathbf{S}^{\mathbf{x}}\mathbf{x} - \mathbf{s}^{\mathbf{x}})^T \lambda^{\mathbf{x}}\|_2^2 + \|(\mathbf{S}^{\mathbf{y}}\mathbf{y} - \mathbf{s}^{\mathbf{y}})^T \lambda^{\mathbf{y}}\|_2^2, \quad (\text{A15})$$

where  $\lambda^{\mathbf{x}} \in \mathbb{R}^{c^{\mathbf{x}}}$  and  $\lambda^{\mathbf{y}} \in \mathbb{R}^{c^{\mathbf{y}}}$  are Lagrangian multipliers.

Optimizing  $E_{\text{IPC}}^{\text{AL}}(\mathbf{y}; \mathbf{x})$  yields the solution to the constrained system:

$$\mathbf{y}^{n+1}; \mathbf{x}^{n+1} = \arg \min_{\mathbf{y}; \mathbf{x}} E_{\text{IPC}}(\mathbf{y}; \mathbf{x}), \quad (\text{A16})$$

$$\text{s.t. } \mathbf{S}^{\mathbf{x}}\mathbf{x} = \mathbf{s}^{\mathbf{x}} \quad \text{and} \quad \mathbf{S}^{\mathbf{y}}\mathbf{y} = \mathbf{s}^{\mathbf{y}}. \quad (\text{A17})$$

## A.6 Guides on Setting the Materials Parameters for Gel Pads

We provide guidance on properly choosing the parameters of the gel pads, which plays a vital role in vision-based tactile sensing. Emperically, the pad’s Poissons ratio  $\nu$  is typically in  $[0.3, 0.45]$ , within which the response variation is subtle. The Youngs modulus typically falls in  $[0.01, 10]\text{MPa}$ , and the response remains almost constant within the same order of magnitude. For simulations without measuring the parameters of the material, our practice is to start within the typical range and carefully adjust them by inspecting the simulation scenes.

While the simulation is not overly sensitive to these parameters, we still recommend measuring these parameters using standard techniques. For example, the sliding experiments for friction coefficients [62] or tensile tests for elastic moduli [18, 27, 33].

## B Additional Related Work

### B.1 Robot Tactile Sensors

Tactile sensing plays a fundamental role in precise manipulation, as established by neuroscientific studies [54, 29, 28, 3]. This understanding has driven the development of artificial tactile sensing systems for robots [44]. Among these, VBTs have gained prominence by offering high-resolution sensing with cost-effectiveness and operational simplicity [61, 53, 37, 35]. While these sensors have advanced robotic manipulation [45, 40, 66], their development remains constrained by the reliance on physical hardware experimentation. **Taccel** addresses this limitation by providing a comprehensive simulation platform to accelerate research and development in tactile robotics.

### B.2 Tactile-Informed Robotic Tasks

Tactile sensing enhances robotic capabilities across three fundamental domains through precise contact interaction measurements:

**Perception** Tactile feedback enables sophisticated object understanding through contact-based sensing. Applications include shear and slip detection [63, 11], object classification and pose estimation [34, 58, 50, 2], material property inference [19, 24], and interaction reconstruction [50, 60, 57]. These perceptual capabilities form the foundation for advanced manipulation algorithms.

**Grasping** Stable grasping requires precise control of contact forces to balance external loads [15, 48]. Tactile sensing provides direct force-torque feedback essential for diverse grasping strategies [39, 36, 59]. This tactile information complements vision-based approaches by enabling fine-grained contact monitoring and in-hand adjustments [6, 5].

**Manipulation** Tactile feedback enables complex manipulation beyond basic pick-and-place operations. Applications include precision tasks like peg insertion [8], object pivoting [21], and articulated object manipulation [67, 3]. Systems such as Tac-Man [66] and DoorBot [52] demonstrate how tactile sensing guides contact geometry understanding and articulation control. This sensing modality is particularly crucial for high-frequency object tracking during dexterous manipulation.

We validate **Taccel**’s capabilities through three representative applications: (i) multi-platform robotic grasping with both rigid and soft objects, (ii) object classification using purely synthetic

training data with strong real-world transfer, and (iii) articulated object manipulation including drawers, cabinets, and bolt-nut assembly tasks, extending the Tac-Man framework [66].

## C Additional Tests on Simulation Speed

### C.1 VRAM Usage

**Tab. A1** report the VRAM occupancy of the low-resolution peg-insertion test (Sec. 5.4). The GPU memory (VRAM) scales linearly with the number of environments, enabling efficient scaling until memory saturation.

Table A1: **VRAM Usage of Tacce1 in the low-resolution peg-insertion test.**

# Envs	1	16	256	1024	4096
VRAM / GiB	4.3	4.3	7.2	16.7	54.0

### C.2 Simulation Speed on Various GPUs

**Tab. A2** reports the simulation speed of the low-resolution peg-insertion test on various GPUs (Sec. 5.4). While HPC GPUs with high FP32:FP64 ratio (1:2) is ideal for optimal performance, desktop GPUs like the 3090 (FP32:FP64 ratio = 1:64) also deliver scalable performance, achieving real-time performance (50FPS) with around 100 environments.

Table A2: **Simulation speed of Tacce1 in the low-resolution peg-insertion task on various GPUs.**

GPU	H100	RTX 4090	RTX 3090	RTX 3090	RTX 3090
# Envs	256	256	256	128	64
Total FPS $\uparrow$	185.52	129.12	121.38	74.81	43.40

## D Experiments Details

### D.1 Real-world Data Collection for Tactile Signal Evaluation

For real-world data collection, our experimental setup consisted of a calibrated GelSight-type sensor, mounted on a vertical rack for precise movement control, as shown in Fig. 3(a). A white mount is fixed in the center of the platform to hold the test objects.

We 3D-print the 18 objects from a standard tactile shape testing dataset [16] at 0.2mm layer height. For each object, we first fix it on the mount, slide the sensor along the rack to press its gel pad on the object, and put a 500 g weight on the sensor to ensure an appropriate amount of gel deformation. We record the RGB tactile patterns five times and compute their mean image to partially remove the noise. We then replicated this pressing sequence in **Tacce1** using a high-resolution soft body gel pad (maximal cell volume  $V_{\max} \approx 10^{-12}\text{m}^3$ ) to ensure signal fidelity. Sec. 5.2 compares the similarity between the real-world and simulated tactile signals.

### D.2 Multi-environment Simulation Test

Our multi-environment simulation test involves three scripted tasks. First, we implemented a peg-insertion task adapted from SAPIEN-IPC [8], where two gel pads (139 nodes and 317 cells each) follow a scripted trajectory. The trajectory involves 200 simulation steps, where the sensors squeeze the peg and manipulate it around the hole. Next, we scaled this task to a higher resolution (1,533 nodes and 5,360 cells each), which results in a larger system to solve. Finally, to demonstrate **Tacce1**’s potential for advanced robotics research, we created a scripted grasping task with a customized five-fingered dexterous hand equipped with 17 gel pads covering the entire hand, totaling 5,157 nodes and 14,311 cells. The trajectory also involves 200 simulation steps, with the hand approaches, grasps, lifts, maneuvers, and releases a stiff cylinder.

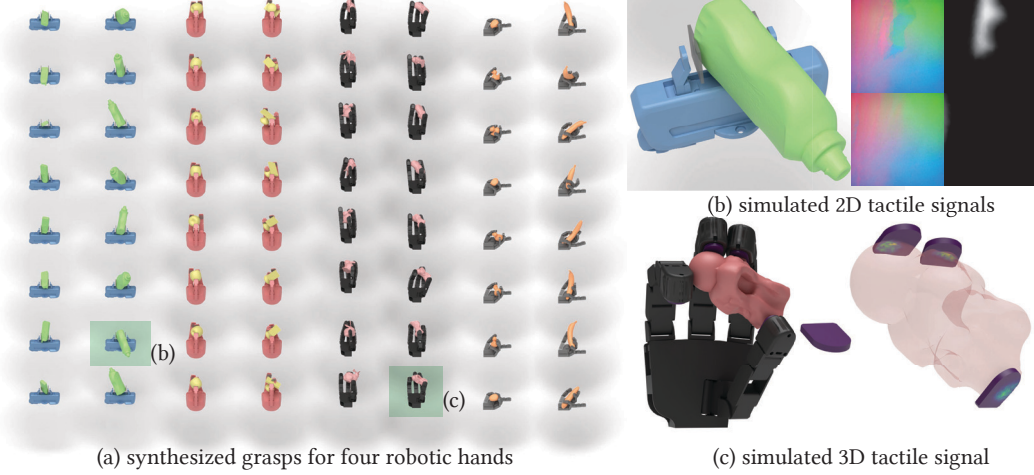


Figure A1: **Examples of the synthesized grasps and the simulated tactile signals in 2D and 3D.** (a) Diverse grasps are synthesized and simulated for different objects and robotic hands. (b) An example of the 2D tactile signals generated for a grasp with the Panda Hand. (c) An example of the 3D tactile signals generated for a grasp with the Allegro Hand with Digit sensors.

### D.3 Data Collection and Training Details for Learning Object Classification

In the object classification sim-to-real experiment, we collect tactile signals of grasping 10 3D-printed mechanical parts in both simulation and the real world. We also train an object classification model with the simulated data and evaluate it on the real data.

For each object, we simulated 200 grasp trials using a parallel gripper with randomized grasping poses. The gripper closes towards the object until the depth deformation exceeds a threshold randomly sampled from  $\tau_d \sim U[0.5, 1.5]\text{mm}$ , and with each grasp yields 2 depth map samples after grasping. To ensure the granularity of the depth map, high-resolution gel pad models  $\max V \geq 10^{-12}\text{m}^3$ . The depth maps  $d_{(u,v)}$  extracted from these simulated tactile signals yielded approximately 4,000 training samples. We split the samples into a training set (85%) and a validation set (15%). The former is used to train the classification model, supervised by the NLL Loss, using the Adam optimizer with a learning rate  $1e-4$  for 100 epochs. During training, we apply an exponential LR scheduler.

For real-world validation, we collected tactile signals using a RobotiQ-2F85 parallel gripper equipped with GelSight-type sensors. The test objects were 3D printed at 0.2mm layer height to maintain high geometric fidelity. We gathered 8 grasps (16 depth maps) per object.

### D.4 Dexterous Grasping

Our key modification to the DFC algorithm promotes perpendicular contact between gel pads and object surfaces, optimizing for downstream tasks that rely on tactile perception.

Following Liu *et al.* [39], we synthesize grasping poses in the robot’s joint space  $q \in \mathcal{Q}$  relative to the object frame by minimizing a modified Gibbs energy:

$$E(O, q, T) = E_{\text{DFC}} + \lambda_{\text{contact}} E_{\text{contact}}(O, q, T). \quad (\text{A18})$$

The force-closure term  $E_{\text{DFC}}$  maintains its original formulation [39], with added constraints to ensure gel pad penetration depth  $\epsilon$  (typically 0.5 mm). We introduce a new contact term  $E_{\text{contact}}$  that aligns the normals of gel pad contacts  $c_i \in \partial^+ \mathcal{G}_i$  with their corresponding object surface normals  $o_i = \arg \min_{o \in \partial O} \|o - c_i\|$ :

$$E_{\text{contact}}(O, q, T) = 1 - \langle c_i^\perp, o_i^\perp \rangle, \quad (\text{A19})$$

where  $(\cdot)^\perp$  represents the surface normals of the gel pad and object surfaces.

**Fig. A1** shows the examples of the synthesized grasps and the simulated tactile signals in both 2D and 3D.

## D.5 Articulated Object Manipulation with Tac-Man

In the Tac-Man framework [66], the system alternates between *execution* and *recovery* phases. During execution, the system performs coarse manipulation actions (*e.g.*, pulling backward) to gradually move the articulated object part. When the actual motion deviates from intended trajectories due to articulation constraints, the gel pad deforms, creating contact deviation reflected in marker flow magnitudes. Once these flows exceed threshold  $\delta_0$ , the system enters recovery mode to restore stable contact by reducing deviation, before resuming execution. This execution-recovery cycle typically requires tens of iterations to finish manipulation.

For our sim-real comparison, we manually created URDF models for the microwave oven and the drawer to match real-world geometries and kinematics, maintaining identical initial grasping poses across simulation and physical setups. Following Tac-Man’s implementation, we set  $\delta_0 = 0.4$  mm and  $\alpha = 0.6$ .