

GraspADMM: Improving Dexterous Grasp Synthesis via ADMM Optimization

Liangwang Ruan^{1,2*}, Jiayi Chen^{1,2*}, He Wang^{1,2,3}, Baoquan Chen^{4†}

Abstract—Synthesizing high-quality dexterous grasps is a fundamental challenge in robot manipulation, requiring adherence to diversity, kinematic feasibility (valid hand-object contact without penetration), and dynamic stability (secure multi-contact forces). The recent framework Dexonomy successfully ensures broad grasp diversity through dense sampling and improves kinematic feasibility via a simulator-based refinement method that excels at resolving exact collisions. However, its reliance on fixed contact points restricts the hand’s reachability and prevents the optimization of grasp metrics for dynamic stability. Conversely, purely gradient-based optimizers can maximize dynamic stability but rely on simplified contact approximations that inevitably cause physical penetrations. To bridge this gap, we propose GraspADMM, a novel grasp synthesis framework that preserves sampling-based diversity while improving kinematic feasibility and dynamic stability. By formulating the refinement stage using the Alternating Direction Method of Multipliers (ADMM), we decouple the target contact points on the object from the actual contact locations on the hand. This decomposition allows the pipeline to alternate between updating the target object points to directly maximize dynamic grasp metrics, and adjusting the hand pose to physically reach these targets while strictly respecting collision boundaries. Extensive experiments demonstrate that GraspADMM significantly outperforms state-of-the-art baselines, achieving a nearly 15% absolute improvement in grasp success rate for type-unaware synthesis and roughly a 100% relative improvement in type-aware synthesis. Furthermore, our approach maintains robust, physically plausible grasp generation even under extreme low-friction conditions.

I. INTRODUCTION

Dexterous grasping is essential for general-purpose robot manipulation. With high degrees of freedom (DoF, typically > 20), dexterous hands can adapt to diverse object shapes. However, this flexibility makes grasp planning highly complex. While data-driven methods have succeeded for parallel grippers by scaling up dataset sizes [1], [2], their extension to dexterous hands is limited by a lack of high-quality datasets. Building these datasets requires solving four fundamental challenges: (1) *diversity*, ensuring broad coverage of objects, hand poses, and grasp types; (2) *kinematic feasibility*, achieving hand-object contact without penetration; (3) *dynamic stability*, generating sufficient contact forces to balance the object’s gravity; and (4) *speed*, achieving fast synthesis.

To address these requirements, analytic methods that optimize metrics like force closure [3] have been explored as

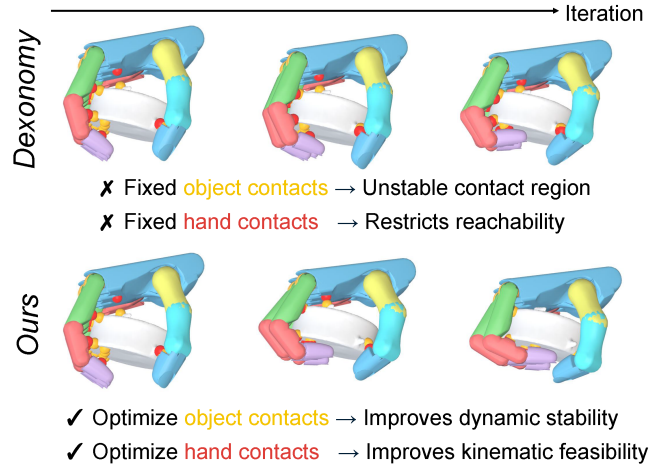


Fig. 1: **Overview.** From the same initialization, our GraspADMM framework generates robust grasps by optimizing contact points on the hand (red) and the object (yellow), while Dexonomy [13] simply fixes these points.

a promising direction. Frameworks like GraspIt! [4] used sampling to handle these non-differentiable metrics, but were highly inefficient for high-DoF hands. Recent works [5]–[12] improved efficiency by introducing differentiable grasp metrics for gradient-based optimization. However, they often focus only on specific grasp types (e.g., fingertips) and struggle with high penetration depths, leading to poor diversity and kinematic feasibility.

Recently, Dexonomy [13] proposed a hybrid approach that improves most of these four aspects compared to prior works. It ensures broad grasp diversity through human-annotated templates and dense sampling, and achieves competitive synthesis speeds. Notably, Dexonomy introduces simulator-based refinement as a promising direction for kinematic feasibility. By utilizing transposed Jacobian control—acting as virtual springs—to pull the hand toward target object points within the forward step of a physics simulator, it robustly resolves exact collisions, demonstrating much better kinematic feasibility than previous gradient-based works.

However, Dexonomy does not fully exploit the potential of this simulator-based framework, leaving critical room for improvement. Specifically, its refinement stage fixes both the target object points and the assigned hand contact points (Fig. 1), which introduces two fundamental drawbacks:

- 1) **Fixed object points limit dynamic stability:** The object contact points may be established in suboptimal

¹School of Computer Science, Peking University.

²Galbot.

³Beijing Academy of Artificial Intelligence.

⁴School of Intelligence Science and Technology, Peking University.

* Equal contribution.

† Corresponding author: baoquan@pku.edu.cn

Methodology	Diversity	Kinematic Feasibility	Dynamic Stability	Speed
Sampling-based (e.g., GraspIt! [4])	×	×	~	×
Gradient-based (e.g., MALA [5])	~	×	✓	~
Simulator-based (Dexonomy [13])	✓	~	~	✓
GraspADMM (Ours)	✓	✓	✓	✓

TABLE I: **Comparison of Dexterous Grasp Synthesis Paradigms.** We evaluate existing methodologies based on their ability to satisfy the four core requirements of high-quality grasp generation. (✓: Fully supports; ~: Partially supports / Some work supports; ×: Struggles).

regions, where the resultant contact force cannot hold the object. Although Dexonomy uses grasp metrics as a post-hoc filter, it simply discards unstable grasps and cannot actively improve them.

- 2) **Fixed hand points limit kinematic feasibility:** Without local mobility, the hand’s kinematic constraints may prevent some hand points from reaching the target object points simultaneously, leading to grasp failure.

Overcoming these drawbacks to unify strict collision resolution with active dynamic optimization is highly non-trivial. On one hand, off-the-shelf physics simulators excel at exact collision resolution (kinematics), but they do not natively support incorporating grasp metrics as optimization objectives. On the other hand, purely gradient-based optimizers can directly maximize dynamic grasp metrics, but they struggle to prevent collisions because they often rely on simplified, penetration-prone contact approximations to bypass non-differentiable physics.

In this work, we bridge this gap by replacing the simulator-based refinement stage with an Alternating Direction Method of Multipliers (ADMM) optimization pipeline. Our GraspADMM framework inherits Dexonomy’s strengths—sampling for diversity and simulation for collision resolution—while formally introducing contact mobility to optimize both dynamic stability and kinematic feasibility.

To achieve this, our formulation explicitly separates the optimization variables. We treat the target contact points on the object surface and the actual contact locations on the hand as independent variables, updating them alternately. This separation circumvents the need to backpropagate gradients through complex, discontinuous contact physics, enabling the use of exact mesh-to-mesh collision detection to rigorously maintain kinematic feasibility. Furthermore, this decomposition allows us to independently optimize the target object points to minimize grasp metrics, actively driving dynamic stability. Consequently, this decoupled approach efficiently synthesizes high-quality grasps that strictly satisfy both kinematic and dynamic constraints.

Extensive experiments demonstrate that GraspADMM significantly outperforms state-of-the-art baselines. On a large-scale benchmark of over 5,600 objects, our method achieves a nearly 15% absolute improvement in grasp success rate over Dexonomy [13] for type-unaware synthesis, while maintaining zero penetration depth and similar speed. This performance gap widens in type-aware tasks, where our decoupled ADMM formulation delivers a 20% to 30%

absolute increase (roughly a 100% relative improvement) in success rates across diverse grasp types. Furthermore, under extreme low-friction conditions ($\mu = 0.1$) where prior methods catastrophically fail, GraspADMM still consistently generates diverse and highly stable multi-fingered grasps.

II. RELATED WORK

A. Analytical Grasp Synthesis

In analytical grasp synthesis, algorithms typically assume complete geometric knowledge of an object to generate hand poses, which are subsequently evaluated using analytical grasp quality metrics. The most widely adopted metrics are based on force closure [3], [14]; however, these are traditionally non-differentiable. Consequently, early approaches either sampled a massive volume of grasps to select the highest-scoring candidates [1], [15]–[17] or employed sampling-based optimization techniques such as simulated annealing [4], [18]. While effective for simple parallel grippers and suction cups with low degrees of freedom ($\text{DoF} < 8$), these strategies suffer from the curse of dimensionality and become computationally intractable for dexterous hands, where the DoF often exceeds 20.

To overcome this limitation, recent research has shifted toward differentiable grasp metrics coupled with gradient-based optimization [5]–[12]. These works primarily focus on formulating novel metrics, such as those based on differentiable force closure [5], [8] or differentiable physics simulations [6], [7]. Regarding optimization strategies, the Metropolis-adjusted Langevin Algorithm (MALA) has been utilized to enhance grasp diversity [5], [8]. Furthermore, Dexonomy [13] highlights the critical role of initialization, proposing a hybrid pipeline that combines sampling-based object initialization with gradient-based refinement. Our framework builds upon these gradient-based foundations but replaces standard gradient descent with a novel ADMM-based optimizer to rigorously resolve exact contact boundaries.

An alternative line of research explores contact-region mapping, transferring human grasp demonstrations to novel objects via functional correspondence [19]–[21]. Although this produces semantically meaningful, human-like grasps suitable for high-level task planning, such approaches often sacrifice grasp diversity and struggle to guarantee dynamic stability under varying physical conditions.

B. Alternating Direction Method of Multipliers (ADMM)

The Alternating Direction Method of Multipliers (ADMM) is a well-established technique in convex optimization [22], proving particularly highly effective for constrained and non-smooth optimization problems. By introducing auxiliary variables, ADMM decouples complex, nonlinear optimization problems into multiple simpler sub-problems, significantly accelerating convergence. This property has been widely exploited in computer graphics and physics simulation, such as in Projective Dynamics (PD) [23], [24] and progressive mesh parameterization [25].

ADMM is also highly adept at handling complex constraints, successfully simulating massive non-penetration scenarios like collisions between elastic bodies and hair [26], [27]. Furthermore, by leveraging proximal operators, ADMM offers unique advantages for non-smooth objectives across geometry processing [28], computer animation [29], and image processing [30]. Beyond applied domains, significant theoretical work has focused on improving the ADMM algorithm itself, including Anderson Acceleration [31], stochastic ADMM formulations [32], and extensions to non-convex optimization [33]. Inspired by its success in resolving discontinuous physical constraints in simulation, our work adapts ADMM for dexterous grasp synthesis, utilizing its decoupled nature to satisfy non-smooth exact collision boundaries while simultaneously optimizing dynamic stability.

III. METHOD

A. Pipeline Overview

Our grasp synthesis framework builds upon the foundational pipeline of Dexonomy [13], progressing through four primary stages: template annotation, sampling-based initialization, simulator-based refinement, and template bootstrapping. The core technical innovation of GraspADMM lies within the refinement stage. Specifically, we replace Dexonomy's purely kinematic refinement with our novel ADMM optimization formulation (detailed in Sec. III-C), which enables the simultaneous optimization of dynamic grasp metrics and exact collision boundaries.

To contextualize this contribution, we first briefly review the preliminary initialization stages inherited from the Dexonomy architecture. The pipeline begins with human-annotated grasp templates mapped to specific categories within the GRASP taxonomy [34]. These templates include the hand pose $\mathbf{q} \in \mathbb{R}^q$, hand contact points $\mathbf{p}_i^h \in \mathbb{R}^3$, corresponding inward-pointing normals $\mathbf{n}_i^h \in \mathbb{R}^3$, and the link assignments for each contact point, where $i \in \{1, 2, \dots, m\}$.

During the initialization stage, the method establishes an initial contact alignment between the object and the grasp template through a hybrid strategy of dense sampling and massively parallel GPU optimization. Specifically, the system randomly selects a grasp template, a hand contact point, and an object surface point. It initializes the object pose by matching the contact locations while aligning the contact normals in opposing directions. Subsequently, with the hand pose fixed, the method computes the nearest surface

projections for all hand contact points and optimizes the object's rigid transformation by minimizing the geometric discrepancy between the hand and object contact points. After this optimization, the results undergo sequential post-filtering to remove severe physical penetrations.

In the ADMM optimization stage, we fix the object pose and optimize the hand pose $\mathbf{q}$, along with auxiliary contact variables, to minimize the dynamic grasp quality metric. Upon convergence, we filter out suboptimal solutions that exhibit poor grasp metrics, unresolved geometric penetrations, or a failure to meet the topological contact requirements defined by the template. Successful grasps are then passed to a rigorous physics simulator evaluation stage and subsequently used to update the templates. For comprehensive details regarding the non-optimization components of this pipeline, please refer to Dexonomy [13].

B. Grasp Quality Metric

While Dexonomy only use grasp quality metric for post-hoc filter, our work use it as the optimization objective and thus greatly improve the dynamic stability. In this section, we summarize the force-closure grasp quality metric utilized in this paper, which aligns with prior works [12], [13]. Note that our algorithm is agnostic to the specific metric and can accommodate other differentiable grasp metrics. For each contact i , let $\mathbf{p}_i \in \mathbb{R}^3$ denote the contact position, $\mathbf{n}_i \in \mathbb{R}^3$ the inward-pointing surface unit normal on the object, and $\mathbf{d}_i \in \mathbb{R}^3$ and $\mathbf{c}_i \in \mathbb{R}^3$ be two orthogonal unit tangent vectors satisfying $\mathbf{n}_i = \mathbf{d}_i \times \mathbf{c}_i$. The Coulomb friction cone $\mathcal{F}_i$ and the contact Jacobian $\mathbf{J}_{o,i}$ for the object at contact i are defined as follows:

$$\mathcal{F}_i = \{\mathbf{x}_i \in \mathbb{R}^3 \mid 0 \leq x_{i,1} \leq 1, x_{i,2}^2 + x_{i,3}^2 \leq \mu^2 x_{i,1}^2\}, \quad (1)$$

$$\mathbf{J}_{o,i}^T = \begin{bmatrix} \mathbf{n}_i & \mathbf{d}_i & \mathbf{c}_i \\ \mathbf{p}_i \times \mathbf{n}_i & \mathbf{p}_i \times \mathbf{d}_i & \mathbf{p}_i \times \mathbf{c}_i \end{bmatrix} \in \mathbb{R}^{6 \times 3}, \quad (2)$$

where μ is the friction coefficient. The friction cone $\mathcal{F}_i$ represents all physically feasible contact forces at contact i , and $\mathbf{J}_{o,i}$ maps a local contact force $\mathbf{x}_i$ to a global object wrench $\mathbf{w}_i = \mathbf{J}_{o,i}^T \mathbf{x}_i$.

To balance an external wrench $\mathbf{g} \in \mathbb{R}^6$ (e.g., object gravity), the optimal contact forces $\{\mathbf{f}_i\}_{i=1}^m$ are obtained by solving the following quadratic program (QP):

$$(\mathbf{f}_1, \dots, \mathbf{f}_m) = \arg \min_{(\mathbf{x}_1, \dots, \mathbf{x}_m)} \left\| \sum_{i=1}^m \mathbf{J}_{o,i}^T \mathbf{x}_i - \mathbf{g} \right\|^2, \quad (3)$$

$$\text{s.t. } \mathbf{x}_i \in \mathcal{F}_i, \quad i \in \{1, \dots, m\}, \quad (4)$$

$$\sum_{i=1}^m x_{i,1} \geq \lambda_{\min}, \quad (5)$$

where $\lambda_{\min}$ is a hyperparameter enforcing a minimum total normal force to prevent the solver from returning trivial zero-force solutions. To reduce computational complexity, the nonlinear friction cone $\mathcal{F}_i$ is approximated by a polyhedral pyramid, converting the problem into a linearly-constrained QP that can be efficiently solved using Clarabel [35].

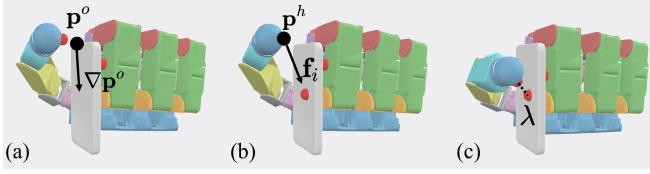


Fig. 2: **ADMM Optimization Pipeline.** (a) Update target object contact points $\mathbf{p}^o$ via gradient descent to maximize dynamic stability. (b) Update hand pose $\mathbf{q}$ and points $\mathbf{p}^h$ via forward-simulated transposed Jacobian control to satisfy kinematic feasibility. (c) Update the dual variable λ .

Finally, the dynamic grasp quality metric e is defined as the residual wrench magnitude:

$$e = \left\| \sum_{i=1}^m \mathbf{J}_{o,i}^T \mathbf{f}_i - \mathbf{g} \right\|^2, \quad (6)$$

where a lower e indicates a more stable and robust grasp. Following [36], we set $\mathbf{g} = \mathbf{0}$ in Eq. 3 rather than systematically testing six orthogonal gravity vectors as done in [12]. This simplification reduces the required QP solves from six to one, significantly improving computational efficiency without compromising grasp quality.

C. ADMM Optimization

Starting from the initialized pose $\mathbf{q}$, the optimization problem in this stage is formally defined as:

$$\min_{\mathbf{q} \in \mathcal{C} \cap \mathcal{T}} e(\mathbf{p}^o, \mathbf{n}^o). \quad (7)$$

Here, e is the QP grasp quality metric from Eq. 6, formulated as a function of the target object contact points $\mathbf{p}^o = \{\mathbf{p}_i^o\}$ and their corresponding normals $\mathbf{n}^o = \{\mathbf{n}_i^o\}$. The constraint $\mathcal{C}$ denotes the strictly penetration-free kinematic workspace of the hand, and $\mathcal{T}$ represents the hand-object contact constraint enforcing $\mathbf{p}_i^h = \mathbf{p}_i^o$ (i.e., the hand contact points must exactly touch the target object points).

Instead of solving this highly nonlinear constrained optimization directly, we employ ADMM to decouple the hard contact constraint $\mathcal{T}$. This yields the following augmented Lagrangian:

$$\begin{aligned} \mathcal{L}(\mathbf{p}^o, \mathbf{p}^h, \lambda) &= e(\mathbf{p}^o, \mathbf{n}^o) + \rho \lambda^T (\mathbf{p}^h - \mathbf{p}^o) + \frac{\rho}{2} \|\mathbf{p}^h - \mathbf{p}^o\|^2 \\ &= e(\mathbf{p}^o, \mathbf{n}^o) + \frac{\rho}{2} \|\mathbf{p}^h - \mathbf{p}^o + \lambda\|^2 - \frac{\rho}{2} \|\lambda\|^2, \end{aligned} \quad (8)$$

where λ is the scaled dual variable (Lagrange multiplier) and ρ is the penalty stiffness parameter. The original problem in Eq. 7 is thus reformulated into a saddle-point problem:

$$\min_{\mathbf{q} \in \mathcal{C}, \mathbf{p}^h, \mathbf{p}^o} \max_{\lambda} \mathcal{L}(\mathbf{p}^o, \mathbf{p}^h, \lambda). \quad (9)$$

We solve Eq. 9 through the following three-step alternating minimization process:

a) 1. Fix $\mathbf{q}, \mathbf{p}^h, \lambda$, update $\mathbf{p}^o$: The subproblem for the target object points is:

$$\min_{\mathbf{p}^o} e(\mathbf{p}^o, \mathbf{n}^o) + \frac{\rho}{2} \|\mathbf{p}^h - \mathbf{p}^o + \lambda\|^2, \quad (10)$$

where $\mathbf{p}^o$ is strictly constrained to the object surface. We use a one-step gradient descent to optimize Eq. 10 and project $\mathbf{p}^o$ back onto the object surface after each update. To compute the gradient of the surface normal $\mathbf{n}^o$ with respect to $\mathbf{p}^o$, we utilize barycentric interpolation across the local mesh triangles, following the methodology in [7].

b) 2. Fix $\mathbf{p}^o, \lambda$, update $\mathbf{q}, \mathbf{p}^h$: The subproblem for the hand kinematics is:

$$\min_{\mathbf{q} \in \mathcal{C}, \mathbf{p}^h} \frac{\rho}{2} \|\mathbf{p}^h - \mathbf{p}^o + \lambda\|^2. \quad (11)$$

This step is mathematically equivalent to minimizing the Euclidean distance between the hand points $\mathbf{p}^h$ and the shifted target points $(\mathbf{p}^o - \lambda)$ while strictly maintaining an interpenetration-free hand state ($\mathbf{q} \in \mathcal{C}$). To robustly optimize this non-convex spatial target, we utilize forward physics simulation in MuJoCo. At each hand contact point $\mathbf{p}_i^h$, we apply a virtual spring force $\mathbf{f}_i$ to bridge the distance gap. This force is mapped to the hand's joint torques τ via transposed Jacobian control:

$$\mathbf{f}_i = k_f (\mathbf{p}^o - \lambda - \mathbf{p}_i^h), \quad \tau = \sum_{i=1}^m \mathbf{J}_{h,i}^T \mathbf{f}_i, \quad (12)$$

where k_f is the virtual spring stiffness, and $\mathbf{J}_{h,i}^T \in \mathbb{R}^{q \times 3}$ is the transpose of the hand's contact Jacobian. By stepping the simulation forward, MuJoCo automatically and robustly resolves all physical collisions. Since we allow $\mathbf{p}_i^h$ to slide freely within its assigned hand link, we project the target point $\mathbf{p}_i^o$ onto the surface of the corresponding hand link to establish the updated $\mathbf{p}_i^h$ at each timestep.

c) 3. Fix $\mathbf{q}, \mathbf{p}^h, \mathbf{p}^o$, update λ : We update the dual variable using the standard ADMM dual ascent step:

$$\lambda \leftarrow \lambda + \mathbf{p}^h - \mathbf{p}^o. \quad (13)$$

Due to the extreme non-linearity of the constraints $\mathcal{C}$ and $\mathcal{T}$, the algorithm is not strictly guaranteed to converge to a perfectly zero-error feasible solution ($\mathbf{p}^h = \mathbf{p}^o$). If the residual error accumulates across multiple steps, the optimization may diverge. To ensure stability, we implement a dual-variable annealing strategy: if the magnitude $\|\lambda_i\| > 2 \times 10^{-2}$ for any contact point i during the optimization loop, we manually reset $\lambda_i = 0$ immediately after Eq. 13.

The decoupled nature of our ADMM formulation can be intuitively understood by examining the extremes of the penalty parameter ρ . If $\rho = 0$, the Lagrangian in Eq. 8 reduces entirely to the grasp metric e . In this scenario, the algorithm simply optimizes the target points on the object surface without considering the hand's kinematic reachability, inevitably resulting in dynamic targets the hand physically cannot grasp. Conversely, as $\rho \rightarrow \infty$ (or if we set $e = 0$), the Lagrangian is dominated by the geometric constraint $\mathbf{p}^h = \mathbf{p}^o$. This reduces the system to a purely

Algorithm 1 GraspADMM Optimization Pipeline

Require: Initial pose $\mathbf{q}_0$ (from Dexonomy initialization);
Object mesh $\mathcal{M}$; Template $\mathcal{P}$; Penalty ρ

Ensure: Optimized hand pose $\mathbf{q}^*$

- 1: *// Initialize from Dexonomy prior*
- 2: $\mathbf{q}, \lambda \leftarrow \mathbf{q}_0, 0$
- 3: $\mathbf{p}^o, \mathbf{p}^h \leftarrow \text{InitializeContacts}(\mathbf{q}, \mathcal{P}, \mathcal{M})$
- 4: **while** not converged **and** iter $< K$ **do**
- 5: *// Step 1: Object Contact Update*
- 6: $\mathbf{p}^o \leftarrow \mathbf{p}^o - \alpha \nabla_{\mathbf{p}^o} \mathcal{L}(\mathbf{p}^o, \mathbf{p}^h, \lambda)$ $\triangleright$ Eq. 10
- 7: $\mathbf{p}^o \leftarrow \text{ProjectToSurface}(\mathbf{p}^o, \mathcal{M})$
- 8: *// Step 2: Hand Kinematics Update via Simulation*
- 9: $\mathbf{f}_i \leftarrow k_f(\mathbf{p}_i^o - \lambda_i - \mathbf{p}_i^h), \forall i \in \{1 \dots m\}$ $\triangleright$ Eq. 12
- 10: $\tau \leftarrow \sum_{i=1}^m \mathbf{J}_{h,i}^T \mathbf{f}_i$ $\triangleright$ Transposed Jacobian control
- 11: $\mathbf{q}, \mathbf{p}^h \leftarrow \text{StepMuJoCo}(\mathbf{q}, \tau)$ $\triangleright$ Forward simulation
- 12: *// Step 3: Dual Variable Update & Annealing*
- 13: $\lambda \leftarrow \lambda + (\mathbf{p}^h - \mathbf{p}^o)$ $\triangleright$ Eq. 13
- 14: **for** $i = 1 \dots m$ **do**
- 15: **if** $\|\lambda_i\| > 2 \times 10^{-2}$ **then**
- 16: $\lambda_i \leftarrow 0$ $\triangleright$ Mitigate divergence
- 17: **end if**
- 18: **end for**
- 19: **end while**
- 20: **return** $\mathbf{q}$

kinematic solver—much like the baseline Dexonomy refinement—where the hand perfectly matches the object surface but operates entirely blind to physical stability forces. By selecting an appropriate intermediate value ($\rho = 10^3$ in our experiments), GraspADMM effectively balances these two forces, allowing the algorithm to dynamically hunt for stable force-closure configurations while rigorously respecting the kinematic boundaries of the initialized grasp pose.

Ultimately, by deploying ADMM, we fully decouple the optimization of the abstract grasp metric from the handling of hard collision constraints. This eliminates the need to backpropagate gradients through discontinuous, non-differentiable collision physics. Because each decoupled subproblem is significantly less nonlinear than the joint problem, they can be efficiently solved using specialized optimizers (gradient descent for e , MuJoCo for C). Consequently, as demonstrated in Table II, our framework achieves significantly faster convergence and dramatically fewer penetrations compared to traditional end-to-end optimization methods.

IV. EXPERIMENTS

A. Evaluation Metric

Our experiment settings and evaluation metrics follow Dexonomy [13]. We make a brief summary below:

- **GSR (%)**: Grasp Success Rate, the percentage of successful grasps out of total input attempts. A grasp

Method	GSR $\uparrow$	OSR $\uparrow$	CDC $\downarrow$	PD $\downarrow$	Div $\downarrow$
DexGraspNet [8]	12.1	57.0	7.6	4.9	29.0
FRoGGeR [10]	10.3	55.7	5.0	0.2	27.0
SpringGrasp [11]	7.8	35.4	23.6	16.6	70.2
BODex [12]	49.2	96.6	3.0	0.6	32.5
Dexonomy [13]	60.5	96.5	0.21	0.0	34.2
GraspADMM (Ours)	74.6	97.2	0.17	0.0	25.7

TABLE II: **Type-Unaware Grasp Synthesis Comparison.** Our method greatly outperforms Dexonomy on grasp success rate and diversity.

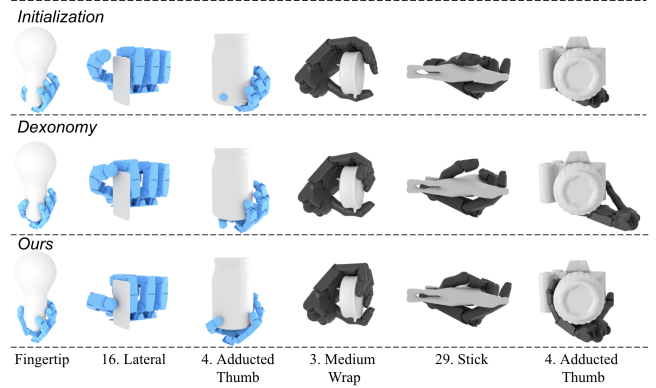


Fig. 3: **Visualization of Optimized Grasps.** Comparison using the Allegro hand (blue) and Shadow hand (black). Given identical initializations, our method synthesizes more physically stable grasps than Dexonomy.

succeeds only if it resists six external forces in MuJoCo. The success criteria for the object pose are 5cm and 15° .

- **OSR (%)**: Object Success Rate, the percentage of objects for which at least one successful grasp is found.
- **CDC (mm)**: Contact Distance Consistency, the delta between the maximum and minimum signed distances across all fingers. A smaller CDC indicates uniform, high-quality contact across all engaged fingers.
- **CLN**: Contact Link Number, the number of hand links whose distance to the object surface is within 2 mm..
- **PD (mm)**: Penetration Depth, encompassing both hand-object inter-penetrations and hand self-collisions.
- **Div (%)**: Diversity, measured by the proportion of the first eigenvalue to the sum of all eigenvalues in the Principal Component Analysis (PCA) of the hand poses.

B. Type-unaware Grasp Synthesis

We adopt the identical experimental setting as Dexonomy [13]: 5,697 assets from DexGraspNet [8], each scaled across 6 geometric sizes ranging from 0.05 to 0.20, are grasped by an Allegro hand using a fingertip style. The baseline’s performance directly comes from the paper of Dexonomy.

As shown in Table II, our method consistently outperforms all baseline approaches across every quality metric. Notably, our method outperform Dexonomy by 14% grasp success rate while keeping the zero penetration property. Moreover,

Grasp Type	GSR (%) $\uparrow$	OSR (%) $\uparrow$	CLN $\uparrow$	Div (%) $\downarrow$
Power	62.8 / 32.9	96.2 / 88.7	12.3 / 9.3	26.4 / 25.2
Intermediate	54.6 / 29.9	96.1 / 89.0	7.4 / 6.4	25.6 / 25.0
Precision	59.2 / 37.5	99.1 / 95.2	5.5 / 4.7	25.1 / 25.0

TABLE III: **Type-Aware Grasp Synthesis Comparison.** Results are presented as (Ours / Dexonomy [13]). We achieve roughly a 100% relative improvement on the grasp success rate.

our grasp is much more diverse, thanks to the flexibility of hand and object contact points. The running time is discussed in Sec. IV-F.

C. Type-Aware Grasp Synthesis

In this section, we evaluate and compare type-aware grasp synthesis with Dexonomy. Because there are around 30 grasp types, it is time-consuming to test on thousands of objects. Therefore, we randomly select 50 objects from DexGraspNet [8] and 50 objects from Objaverse [37]. In our early experiments, we find that this subset yields similar performance to the full object set and is thus representative. Following the protocol in Dexonomy [13], we generate 100 grasp attempts for each object-template pair, and use the same initialization for Dexonomy’s refinement stage and our GraspADMM pipeline.

1) *Quantitative Comparison:* The quantitative statistics for type-aware grasp quality are summarized in Table III. Note that the GSR and OSR metrics for Dexonomy are slightly higher than those reported in the original paper due to minor implementation differences and object randomization. Nevertheless, our method significantly outperforms Dexonomy in both GSR and OSR across all three grasp categories, validating the efficacy of our ADMM optimization. Furthermore, the higher Contact Link Numbers (CLN) demonstrate that GraspADMM achieves richer, multi-link contacts, leading to vastly more secure grasps. Our method generates slightly less diverse poses, which is partially due to more samples passing the simulation validation.

2) *Visual Comparison:* We visually compare our ADMM optimization against the local refinement stage of Dexonomy. As illustrated in Fig. 3, starting from identical initializations, our method yields more visually plausible results for both the Allegro and Shadow hands. While the local refinement stage in Dexonomy fixes contact points on the object and hand links, our method allows adjustments in local frames. Through ADMM optimization, our method encourages lower grasp quality metrics and more contacts, both contribute to more stable grasp poses.

D. Robustness for Low-Friction Objects

The standard simulation validation discussed above assumes tangential and torsional friction coefficients of 0.6 and 0.02, respectively. However, these values often fail to reflect the inherently slippery and unpredictable nature of complex real-world grasping. To rigorously stress-test our approach, we introduce a harder benchmark by eliminating torsional

	GSR $\uparrow$	OSR $\uparrow$	CDC $\downarrow$	PD $\downarrow$
No initialization	27.2	55.1	0.31	0.0
$\rho = 0$	17.8	72.6	10.24	4.95
$\rho = 10^2$	73.5	96.7	0.21	0.0
$\rho = 10^4$	59.3	96.2	0.29	0.0
$\rho \rightarrow \infty$	62.6	94.7	0.04	0.0
Ours ($\rho = 10^3$)	74.6	97.2	0.17	0.0

TABLE IV: **Ablation Study.** A good initialization and a moderate ρ like 10^2 and 10^3 is important.

friction entirely (setting it to 0) and incrementally reducing the tangential friction coefficient from 0.5 down to 0.1.

As shown in Fig. 4, GraspADMM consistently outperforms Dexonomy in both GSR and OSR across all friction regimes. The relative performance gap actually widens as the environment becomes more challenging, underscoring the necessity of explicitly optimizing dynamic stability.

This advantage is especially pronounced under extreme low-friction conditions ($\mu = 0.1$). As visualized in Fig. 5, for the selected object-template pairs, Dexonomy rarely produces more than a single grasp pose that survives the simulation validation. In contrast, our decoupled ADMM optimization reliably generates multiple diverse, physically viable grasp poses from the same pool of initial attempts.

E. Ablation Study

Table IV shows the impact of different modules and parameter choices on the fingertip grasp synthesis experiment for Allegro hand.

1) *Initialization:* Our method uses the global alignment stage from Dexonomy [13] as initialization, which first aligns the hand to the object on one contact point, then optimizes the object pose to try to match all contacts. Without this initialization step and no template annotation, the GSR and OSR of our method significantly drops to 27.18% and 55.10%, showing the importance of initialization.

2) *Choice of ρ :* ρ is the most important hyperparameter in ADMM, as explained in Sec III-C. If $\rho = 0$, the hand pose is not optimized but directly copied from initialization, results in low GSR and OSR. If $\rho \rightarrow \infty$, the contact points on the object are fixed, making our methods perform similarly to Dexonomy [13]. Setting a moderate ρ , like 10^2 and 10^3 , achieves overall high GSR and OSR, but $\rho = 10^4$ is too large, causing performance drops.

3) *Dual variable annealing:* We visualize the importance of dual variable annealing in Fig. 6. The green points represent the object contact points $\mathbf{p}^o$, the red points represent $\mathbf{p}^o - \lambda$. Without annealing, due to the nonzero gap between $\mathbf{p}^h$ and $\mathbf{p}^o$, λ accumulates rapidly across iterations, causing $\mathbf{p}^o - \lambda$ to drastically deviate from $\mathbf{p}^o$. This makes the hand become unstable in the second stage of ADMM. After applying annealing, λ is controlled under a certain threshold, letting the hand to stably squeeze as tight as possible.

F. Time Analysis

In terms of optimization speed, both our method and Dexonomy utilize a hybrid CPU (for MuJoCo) and GPU

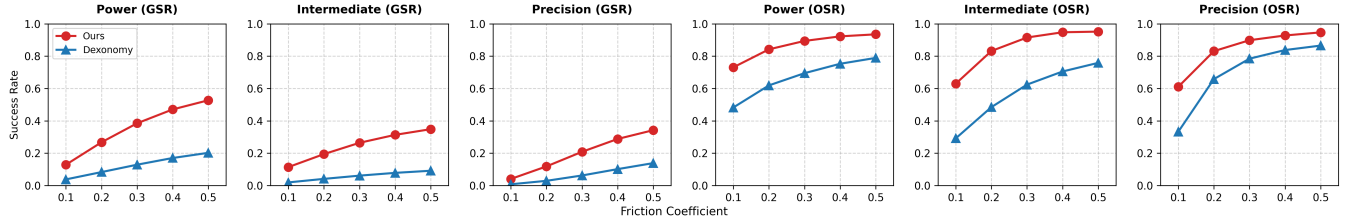


Fig. 4: **Performance on the Hard Benchmark.** Grasp success rate and object success rate across varying tangential friction coefficients. Our method consistently outperforms Dexonomy.

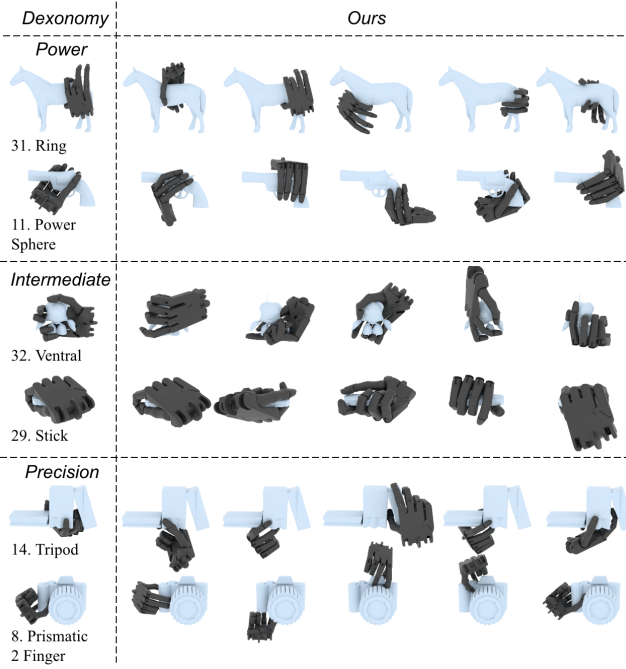


Fig. 5: **Robustness under Extreme Low Friction ($\mu = 0.1$).** While Dexonomy struggles to find more than a single stable grasp out of 100 attempts for the shown objects, GraspADMM reliably generates a diverse set of stable grasps.

pipeline, whereas the other baselines rely primarily on GPU optimization. On a single NVIDIA RTX 4090 GPU, DexGraspNet [8], FRoGGeR [10], and SpringGrasp [11] average fewer than 3 samples per second (s/s). BODex [12] reaches approximately 40 s/s. Utilizing a single AMD EPYC 7543 2.8GHz CPU, Dexonomy processes around 24 s/s. Our method requires slightly more computation time due to the iterative ADMM optimization, processing approximately 19 s/s. However, while BODex boasts the highest throughput, it is strictly limited to fingertip-style grasping, whereas our framework inherently supports the full taxonomy of diverse grasp types. Furthermore, because our method achieves a significantly higher GSR than Dexonomy, the actual yield of successful, usable grasp poses generated per second is highly comparable. Future extensions, such as scaling CPU resources or transitioning to GPU-native simulators like NVIDIA Newton [38], could further accelerate our pipeline.

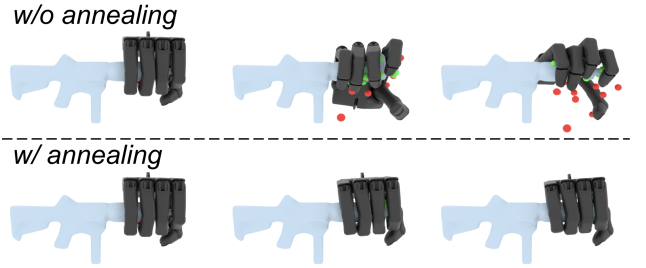


Fig. 6: **Effect of Dual Variable Annealing.** Green: target object contacts p^o . Red: dual-shifted targets $p^o - \lambda$. Annealing bounds the accumulation of λ , preventing the shifted targets from diverging and ensuring stable optimization.

V. CONCLUSION & LIMITATION

In this work, we presented GraspADMM, a novel dexterous grasp synthesis framework that improves the kinematic feasibility and dynamic stability. By formulating the refinement stage via the Alternating Direction Method of Multipliers (ADMM), our method successfully decouples the object contact points from hand contact points, enabling the efficient minimization of grasp metrics while enforcing penetration-free kinematic constraints. Extensive experiments demonstrate that GraspADMM achieves state-of-the-art success rates across diverse grasp types and object geometries, maintaining robust even under extreme low-friction conditions (e.g., $\mu = 0.1$).

Nevertheless, our work faces several limitations. First, currently we use CPU-based simulator and multiprocessing which causes performance drop. A full GPU pipeline could potentially improve efficiency and throughput. Second, our work has not yet generated a large-scale dataset for learning-based methods. Last but not least, our work focuses on static grasping in a non-clustering scene. A promising direction for future research is to incorporate trajectory generation for dynamic grasping, and extend to more complex clustered scenes.

REFERENCES

- [1] H.-S. Fang, C. Wang, M. Gou, and C. Lu, “Graspnet-1billion: A large-scale benchmark for general object grasping,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 11 444–11 453.

- [2] H.-S. Fang, C. Wang, H. Fang, M. Gou, J. Liu, H. Yan, W. Liu, Y. Xie, and C. Lu, "Anygrasp: Robust and efficient grasp perception in spatial and temporal domains," *IEEE Transactions on Robotics*, vol. 39, no. 5, pp. 3929–3945, 2023.
- [3] C. Ferrari, J. F. Canny *et al.*, "Planning optimal grasps." in *ICRA*, vol. 3, no. 4, 1992, p. 6.
- [4] A. T. Miller and P. K. Allen, "Graspt! a versatile simulator for robotic grasping," *IEEE Robotics & Automation Magazine*, vol. 11, no. 4, pp. 110–122, 2004.
- [5] T. Liu, Z. Liu, Z. Jiao, Y. Zhu, and S.-C. Zhu, "Synthesizing diverse and physically stable grasps with arbitrary hand structures using differentiable force closure estimator," *IEEE Robotics and Automation Letters*, vol. 7, no. 1, pp. 470–477, 2021.
- [6] D. Turpin, L. Wang, E. Heiden, Y.-C. Chen, M. Macklin, S. Tsogkas, S. Dickinson, and A. Garg, "Grasp'd: Differentiable contact-rich grasp synthesis for multi-fingered hands," in *European Conference on Computer Vision*. Springer, 2022, pp. 201–221.
- [7] D. Turpin, T. Zhong, S. Zhang, G. Zhu, E. Heiden, M. Macklin, S. Tsogkas, S. Dickinson, and A. Garg, "Fast-grasp'd: Dexterous multi-finger grasp generation through differentiable simulation," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 8082–8089.
- [8] R. Wang, J. Zhang, J. Chen, Y. Xu, P. Li, T. Liu, and H. Wang, "Dexgraspnet: A large-scale robotic dexterous grasp dataset for general objects based on simulation," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 11 359–11 366.
- [9] J. Chen, Y. Chen, J. Zhang, and H. Wang, "Task-oriented dexterous grasp synthesis via differentiable grasp wrench boundary estimator," *arXiv preprint arXiv:2309.13586*, 2023.
- [10] A. H. Li, P. Culbertson, J. W. Burdick, and A. D. Ames, "Frogger: Fast robust grasp generation via the min-weight metric," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023, pp. 6809–6816.
- [11] S. Chen, J. Bohg, and C. K. Liu, "Springgrasp: An optimization pipeline for robust and compliant dexterous pre-grasp synthesis," *arXiv preprint arXiv:2404.13532*, 2024.
- [12] J. Chen, Y. Ke, and H. Wang, "Bodex: Scalable and efficient robotic dexterous grasp synthesis using bilevel optimization," *arXiv preprint arXiv:2412.16490*, 2024.
- [13] J. Chen, Y. Ke, L. Peng, and H. Wang, "Dexonomy: Synthesizing all dexterous grasp types in a grasp taxonomy," *Robotics: Science and Systems*, 2025.
- [14] M. A. Roa and R. Suárez, "Grasp quality measures: review and performance," *Autonomous robots*, vol. 38, pp. 65–88, 2015.
- [15] J. Mahler, F. T. Pokorny, B. Hou, M. Roderick, M. Laskey, M. Aubry, K. Kohlhoff, T. Kröger, J. Kuffner, and K. Goldberg, "Dex-net 1.0: A cloud-based network of 3d objects for robust grasp planning using a multi-armed bandit model with correlated rewards," in *2016 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2016, pp. 1957–1964.
- [16] J. Mahler, M. Matl, X. Liu, A. Li, D. Gealy, and K. Goldberg, "Dex-net 3.0: Computing robust vacuum suction grasp targets in point clouds using a new analytic model and deep learning," in *2018 IEEE International Conference on robotics and automation (ICRA)*. IEEE, 2018, pp. 5620–5627.
- [17] H. Cao, H.-S. Fang, W. Liu, and C. Lu, "Suctionnet-1billion: A large-scale benchmark for suction grasping," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 8718–8725, 2021.
- [18] M. Ciocarlie, C. Goldfeder, and P. Allen, "Dexterous grasping via eigengrasps: A low-dimensional approach to a high-complexity problem," in *Robotics: Science and systems manipulation workshop-sensing and adapting to the real world*, 2007.
- [19] L. Yang, K. Li, X. Zhan, F. Wu, A. Xu, L. Liu, and C. Lu, "Oakink: A large-scale knowledge repository for understanding hand-object interaction," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 20 953–20 962.
- [20] W. Wei, P. Wang, S. Wang, Y. Luo, W. Li, D. Li, Y. Huang, and H. Duan, "Learning human-like functional grasping for multi-finger hands from few demonstrations," *IEEE Transactions on Robotics*, 2024.
- [21] R. Wu, T. Zhu, X. Lin, and Y. Sun, "Cross-category functional grasp transfer," *arXiv preprint arXiv:2405.08310*, 2024.
- [22] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [23] T. Liu, A. W. Bargteil, J. F. O'Brien, and L. Kavan, "Fast simulation of mass-spring systems," *ACM Trans. Graph.*, vol. 32, no. 6, Nov. 2013. [Online]. Available: <https://doi.org/10.1145/2508363.2508406>
- [24] R. Narain, M. Overby, and G. E. Brown, "Admm $\supseteq$ projective dynamics: fast simulation of general constitutive models," in *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, ser. SCA '16. Goslar, DEU: Eurographics Association, 2016, p. 21–28.
- [25] L. Liu, C. Ye, R. Ni, and X.-M. Fu, "Progressive parameterizations," *ACM Transactions on Graphics (SIGGRAPH)*, vol. 37, no. 4, 2018.
- [26] G. Daviet, "Simple and scalable frictional contacts for thin nodal objects," *ACM Trans. Graph.*, vol. 39, no. 4, 8 2020. [Online]. Available: <https://doi.org/10.1145/3386569.3392439>
- [27] —, "Interactive hair simulation on the gpu using admm," in *ACM SIGGRAPH 2023 Conference Proceedings*, ser. SIGGRAPH '23. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: <https://doi.org/10.1145/3588432.3591551>
- [28] S. Bouaziz, A. Tagliasacchi, and M. Pauly, "Sparse iterative closest point," in *Proceedings of the Eleventh Eurographics/ACMSIGGRAPH Symposium on Geometry Processing*, ser. SGP '13. Goslar, DEU: Eurographics Association, 2013, p. 113–123. [Online]. Available: <https://doi.org/10.1111/cgf.12178>
- [29] T. Neumann, K. Varanasi, S. Wenger, M. Wacker, M. Magnor, and C. Theobalt, "Sparse localized deformation components," *ACM Trans. Graph.*, vol. 32, no. 6, Nov. 2013. [Online]. Available: <https://doi.org/10.1145/2508363.2508417>
- [30] Y. Hu, D. Zhang, J. Ye, X. Li, and X. He, "Fast and accurate matrix completion via truncated nuclear norm regularization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 9, pp. 2117–2130, 2013.
- [31] J. Zhang, Y. Peng, W. Ouyang, and B. Deng, "Accelerating admm for efficient simulation and optimization," *ACM Trans. Graph.*, vol. 38, no. 6, Nov. 2019. [Online]. Available: <https://doi.org/10.1145/3355089.3356491>
- [32] F. Bian, J. Liang, and X. Zhang, "A stochastic alternating direction method of multipliers for non-smooth and non-convex optimization," *Inverse Problems*, vol. 37, no. 7, p. 075009, jul 2021. [Online]. Available: <https://doi.org/10.1088/1361-6420/ac0966>
- [33] Y. Wang, W. Yin, and J. Zeng, "Global convergence of admm in nonconvex nonsmooth optimization," 2018. [Online]. Available: <https://arxiv.org/abs/1511.06324>
- [34] T. Feix, J. Romero, H.-B. Schmiedmayer, A. M. Dollar, and D. Kragic, "The grasp taxonomy of human grasp types," *IEEE Transactions on human-machine systems*, vol. 46, no. 1, pp. 66–77, 2015.
- [35] P. J. Goulart and Y. Chen, "Clarabel: An interior-point solver for conic programs with quadratic objectives," *arXiv preprint arXiv:2405.12762*, 2024.
- [36] A. Wu, M. Guo, and C. K. Liu, "Learning diverse and physically feasible dexterous grasps with generative model and bilevel optimization," *arXiv preprint arXiv:2207.00195*, 2022.
- [37] M. Deitke, D. Schwenk, J. Salvador, L. Weihs, O. Michel, E. Vander-Bilt, L. Schmidt, K. Ehsani, A. Kembhavi, and A. Farhadi, "Objaverse: A universe of annotated 3d objects," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 13 142–13 153.
- [38] (2025) Newton physics engine documentation. NVIDIA Corporation. Joint development with Google DeepMind and Disney Research; contributed to the Linux Foundation. [Online]. Available: <https://developer.nvidia.com/newton-physics>