

# DG16M: A Large-Scale Dataset for Dual-Arm Grasping with Force-Optimized Grasps

Md Faizal Karim<sup>1\*</sup>, Mohammed Saad Hashmi<sup>1\*</sup>, Shreya Bollimuntha<sup>1</sup>, Mahesh Reddy Tapeti<sup>1</sup>, Gaurav Singh<sup>1</sup>,  
Nagamanikandan Govindan<sup>2</sup>, K Madhava Krishna<sup>1</sup>

\*Equal Contribution, <sup>1</sup>Robotics Research Center, IIIT Hyderabad, <sup>2</sup> IIITDM Kancheepuram

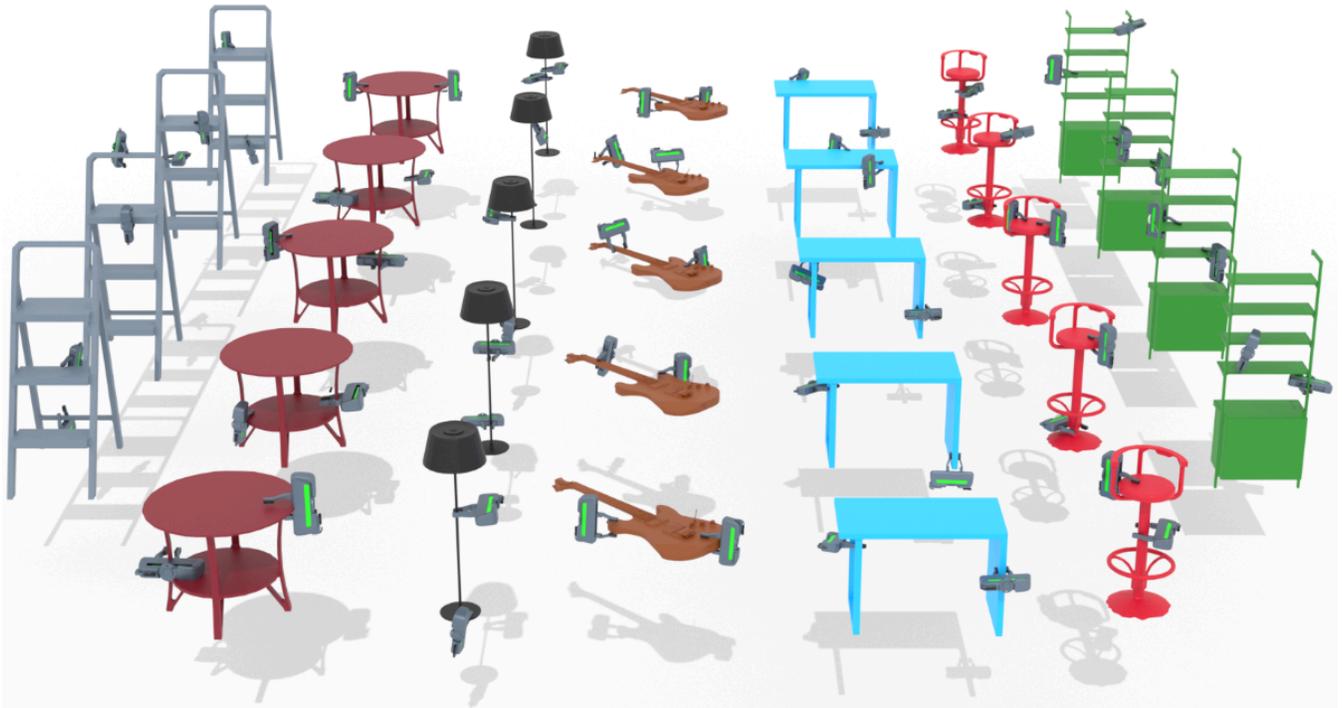


Fig. 1: **Dataset Visualization:** We introduce DG16M, a large-scale dual-arm grasp dataset generated using optimization-based force-closure constraints to ensure stable and physically viable grasps. Our dataset provides high-quality grasp pairs, enabling better generalization for deep learning-based grasp generation models.

**Abstract**—Dual-arm robotic grasping is crucial for handling large objects that require stable and coordinated manipulation. While single-arm grasping has been extensively studied, datasets tailored for dual-arm settings remain scarce. We introduce a large-scale dataset of 16 million dual-arm grasps, evaluated under improved force-closure constraints. Additionally, we develop a benchmark dataset containing 300 objects with approximately 30,000 grasps, evaluated in a physics simulation environment, providing a better grasp quality assessment for dual-arm grasp synthesis methods. Finally, we demonstrate the effectiveness of our dataset by training a Dual-Arm Grasp Classifier network that outperforms the state-of-the-art methods by 15%, achieving higher grasp success rates and improved generalization across objects. Project page: <https://dg16m.github.io/DG-16M/>

## I. INTRODUCTION

Bimanual manipulation is essential in robotic applications involving handling large, heavy, or unwieldy objects that a single robotic arm cannot effectively grasp. Industries such

as automation, logistics, assistive robotics, and assembly frequently require coordinated dual-arm strategies to ensure stable object handling [1]–[5]. A fundamental step in bimanual manipulation is establishing stable grasps on the object that distribute forces evenly, and minimize disturbances across various environments.

Learning-based grasp planning methods have shown significant promise in the single-arm setting, leveraging large-scale datasets to improve generalization across diverse objects [6]–[10]. To sample such datasets, traditional analytical methods are used, that optimize grasp configurations based on object geometry and contact models, employing force-closure and wrench analysis to ensure grasp stability [11], [12]. These approaches identify antipodal contact points, friction cones, and key stability metrics, such as the Ferrari-Canny epsilon metric [12], to rank potential grasps.

Despite the availability of large-scale single-arm grasp datasets, they cannot be directly applied to dual-arm grasp-

Dataset	Grasps per Obj	Grasp Label	Total Obj	Total Grasps	Gripper Type
ACRONYM [13]	2000	6-DoF	8872	17.7M	Single
DA2 [14]	up to 2001	6-DoF	6327	9M	Dual
<b>DG16M (ours)</b>	up to 4000	6-DoF	4132	16M	Dual

TABLE I: Comparison of grasp datasets based on grasp density, labeling method, object count, total grasps, and gripper types supported.

ing. The primary challenge lies in effectively pairing single-arm grasps to create stable dual-arm grasp configurations [5]. Unlike single-arm grasping, where stability is primarily determined by local contact properties, bimanual grasping requires considering additional constraints, such as object size, shape, and the coordination between both grippers. Extending grasp dataset generation to dual-arm settings introduces significant complexity due to these added constraints, making robust dataset construction and learning frameworks a challenging task.

Existing dual-arm grasp dataset, DA2 [14], uses a simplified force-closure formulation that assumes constant normal forces at the contact points and lacks validation through physics simulation as shown in figure 3. As a result, the dataset contains various types of failed grasps. In contrast, single-arm grasp datasets incorporate physics simulations to validate grasps, ensuring a higher degree of reliability [13]. Thus, there is a need for a high-quality dual-arm grasp dataset for the development of data-driven approaches for bimanual grasping.

To address this challenge, we introduce DG16M, a novel dataset that integrates a stricter force-closure formulation which analytically assesses grasp stability while accounting for varying external forces and gripper constraints. Our dataset consists of 4,143 objects of varying shapes and sizes from [14], each with up to 4,000 dual-arm grasp pairs, resulting in a total of approximately 16 million grasps (a small subset is shown in Figure 1). Furthermore, we introduce a benchmark dataset containing 300 objects with approximately 30,000 grasps, verified through physics simulation-based evaluation using Isaac Gym [15].

To summarize our contributions:

- We introduce DG16M, a large-scale dataset of 16 million dual-arm grasps, incorporating an improved force-closure validation to ensure physically stable and reliable grasp assessment. Through both classifier efficacy and analysis of grasp matrix, we demonstrate improvement in dual-arm grasp quality vis-à-vis prior datasets.
- We develop an improved dual-arm grasp classifier leveraging neural descriptor fields, demonstrating that training on grasps vetted by our force-closure formulation yields higher success rates in simulation compared to prior methods.
- We introduce a benchmark dataset of 300 objects with 30,000 simulation-verified grasps, providing a standardized evaluation for dual-arm grasping. By incorporating

physics-based validation, our benchmark overcomes the limitations of force-closure analysis, ensuring a more practical and robust dual-arm grasp stability assessment.

- We also conduct ablation studies to demonstrate that sampling techniques like antipodal sampling alone cannot guarantee stable dual-arm grasp pairs. We show that these grasp pairs achieve stability only when further validated through our force-closure analysis.

## II. RELATED WORK

### A. Grasp generation

Grasp synthesis approaches explore the space of possible grasp to identify stable and feasible grasps. One widely used approach is **approach-based sampling** [16]–[19], which aligns the gripper’s approach vector with the surface normal of the sampled object points. This alignment improves stability by ensuring perpendicular contact. However, it can introduce bias, potentially overlooking viable grasps that do not conform to the sampled surface normal [20].

Another common strategy is **antipodal sampling** [21], which is extensively used in single-arm grasping datasets [11], [13], [14], [21]. This method selects opposing contact points that satisfy force-closure conditions, ensuring a stable grip, particularly for parallel-jaw grippers. Due to its simplicity and effectiveness, antipodal sampling has become a standard technique for generating stable grasps. For dual-arm grasping, we adopt the block antipodal sampling introduced in [14], as the first step in our grasp generation pipeline due to its ability to generate feasible contact points efficiently for large objects. Furthermore, we show in Section (V) that randomly pairing antipodal grasps is insufficient for stable dual-arm grasping, highlighting the necessity of force-closure validation.

### B. Force Closure

Force closure is a fundamental property in robotic grasping, ensuring that an object can be securely held against external disturbances. A grasp achieves force closure if the set of contact forces can counteract any external wrench acting on the object. Traditionally, force closure is computed by constructing a convex hull over discretized friction cones in wrench space under the assumption that the total contact forces sum to one [22], [23]. [24] proposed two relaxations to the traditional formulation, namely, zero friction and equal magnitude of contact forces, to make the formulation differentiable. Even though this formulation is useful for a human hand, the assumptions made pose certain challenges when

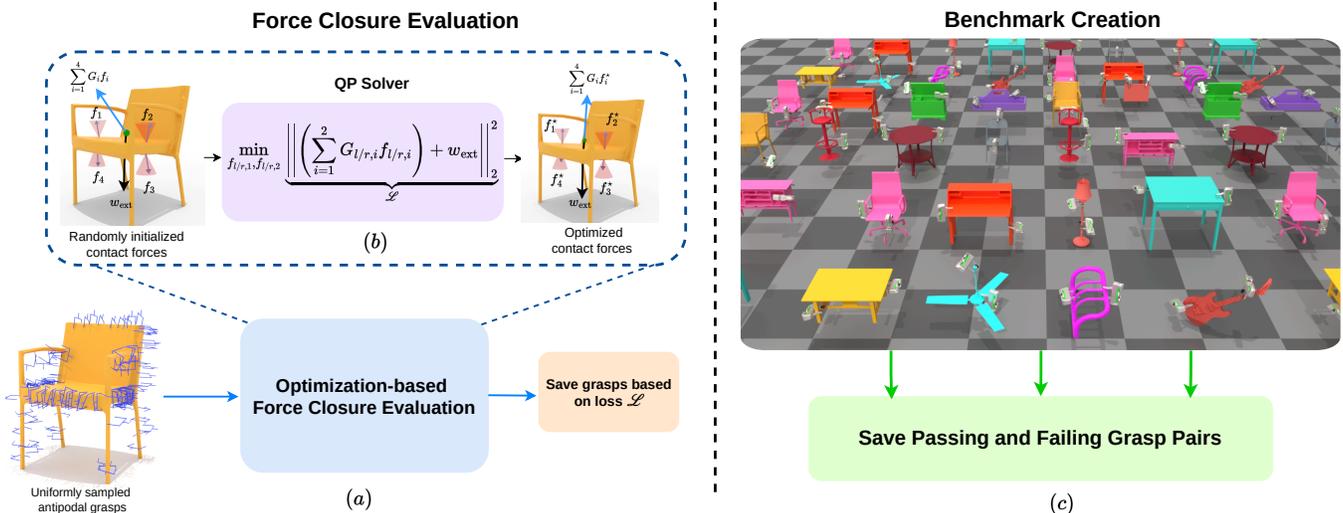


Fig. 2: **Overview of the proposed method:** (a) We start by sampling a large number of antipodal grasps on the object mesh, generate all possible grasp pairs, and apply distance-based pruning. These pairs are evaluated using an Optimizer-based Force Closure Evaluator, which checks if a valid set of contact forces can keep the object in equilibrium under an external wrench. (b) The optimizer solves (2) under constraints (3), (4), enforcing gripper force limits and friction cone constraints. Valid grasp pairs are identified by thresholding the loss values, as shown in (5), ensuring that only stable grasps are retained while unstable ones are discarded. (c) We select a subset of objects and evaluate their grasps in simulation to construct the benchmark dataset. This evaluation provides stability ground truth values for each dual-arm grasp, offering an objective assessment that is independent of predefined grasp quality metrics or assumptions.

adapting it to a dual-gripper setting [14]. PhyGrasp [25] uses an optimization-based force closure formulation to solve for the contact forces. They sample contact points uniformly across the object without any constraints in order to find an affordance map, resulting in some of these grasps being practically infeasible. We propose an optimization-based constrained force closure formulation that addresses the issues with using a relaxed [24] or unconstrained [25] formulation for a dual-gripper setting.

### C. Benchmarking Datasets for Grasping

In the field of robotic grasping, the development of benchmarking datasets has been pivotal for advancing grasp synthesis and manipulation techniques, with most efforts focused on **single-arm** grasping. Several large-scale single-arm datasets, such as [11], [26]–[29], provide RGB-D images of objects along with annotated grasps, either in the form of 6D grasp poses or bounding box markings. In addition, datasets like [8], [13] focus on synthetic point clouds and grasp annotations, expanding the range of objects available for training. While these datasets are extensive and contain a large number of grasps, they are primarily designed for small-sized objects, partly due to the physical constraints of single-arm manipulation. Therefore, they typically use object meshes from YCB [30] and GSO [31] or specific categories of ShapeNet [32] (e.g. bottle, mug, cup, etc.)

Existing dual-arm grasping datasets are limited, with DA2 [14] being the first large-scale dual-arm grasp dataset, containing approximately 9 million grasps across a wide range of large objects selected from the ShapeNet dataset [32] (refer to Table I). While this dataset is extensive, it follows a relatively relaxed set of force closure constraints that do not account for actuator constraints, increasing the gap between

the grasps in the dataset and the grasps which are practically feasible. Conversely, DG16M employs an optimization-based analytical force closure formulation that can account for the actuator constraints and provide optimal contact forces that counteract external wrenches. This results in well-separated positive and negative grasp pairs, enabling classifiers like [33] and [34] to learn effectively. Our dataset can readily be integrated with frameworks such as our previous work DA-VIL [35] to achieve better inferences.

## III. DATASET GENERATION

In this section, we describe the pipeline used to generate grasps for our DG16M dataset as shown in Figure 2. The pipeline consists of two key steps: (1) generating grasp candidates through antipodal sampling, (2) evaluating grasp quality using an optimization-based force closure formulation.

### A. Grasp Sampling

To generate grasp candidates, we employ **antipodal grasp sampling**, a method commonly used for grasp pose generation [11]. The process begins by randomly selecting a point on the object’s surface along with its corresponding normal. A second contact point is then found by tracing a ray within a constrained region around the normal direction, ensuring that the two contact points satisfy antipodal conditions. The allowable range for this ray is determined by a threshold  $\gamma = \tan(\alpha/2)$ , where  $\alpha$  is the cone angle given by  $\alpha = \tan^{-1}(\mu)$ , with  $\mu$  representing the friction coefficient. After identifying both contact points, we verify that the grasp configuration does not result in any collision between the grippers and the object. Any grasp that leads to penetration or overlaps with the object is discarded, ensuring

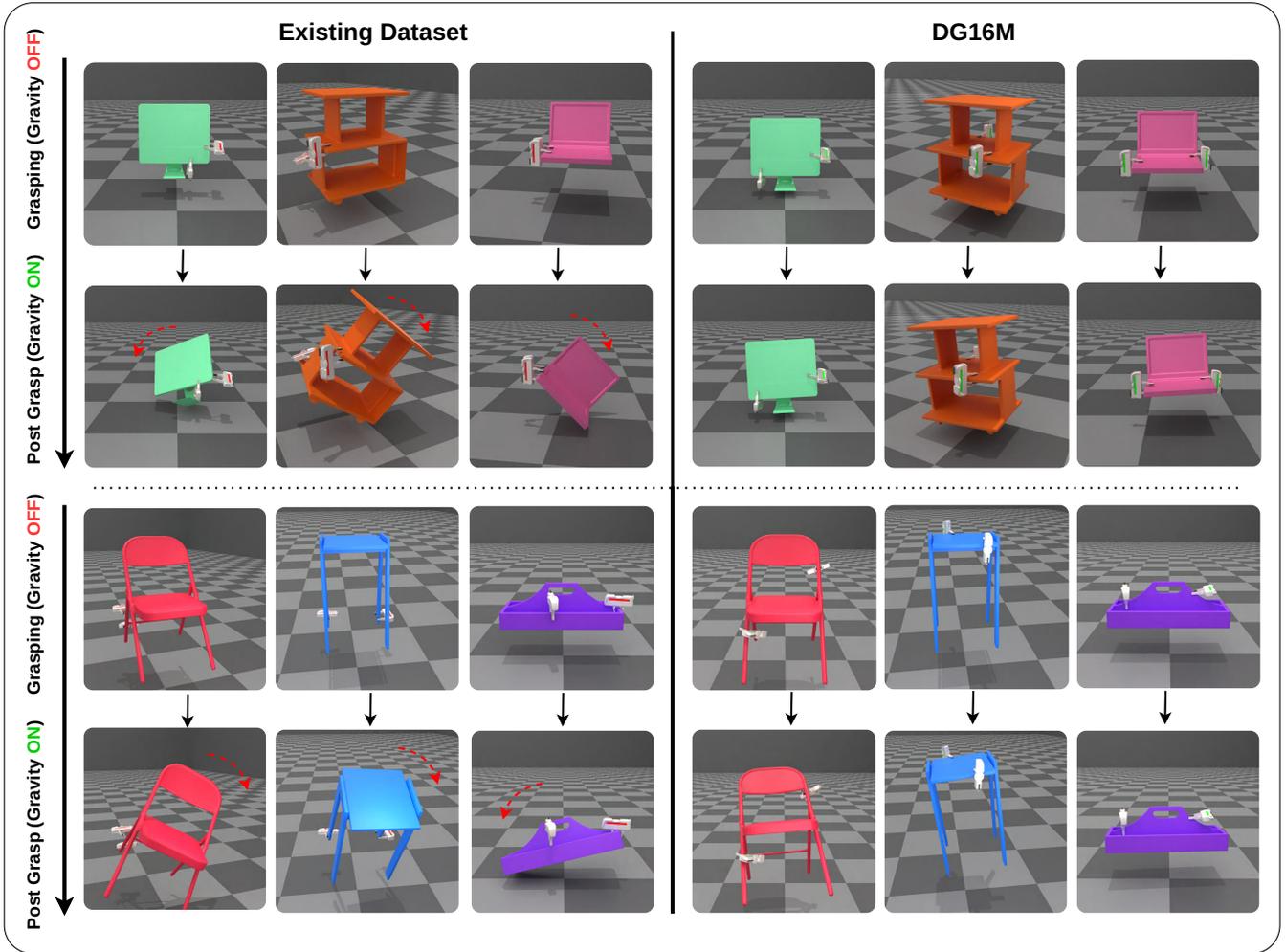


Fig. 3: **Qualitative Comparison between the Existing DA2 Dataset and the Proposed DG16M Dataset** We evaluate grasp stability by initializing the object and floating grippers at the grasp pose with gravity disabled. Once the grippers fully close, gravity is enabled. Grasps from the existing dataset often fail to hold the object in place (indicated by red dotted arrows) or lose contact due to instability, which arises from issues such as grasp pairs being too close, positioned on the same side of the center of mass, or lacking force balance. In contrast, our dataset provides more stable and physically viable grasps, ensuring successful object retention without failures. Floating grippers are shown in white. (*Zoom in for a clearer view*)

that the final sampled grasps are physically valid and free of collisions. We initially sample 500 single-arm grasps per object, which are then combined exhaustively to generate all possible unique combinations of dual-arm grasp pairs. To improve diversity and prevent redundancy, we apply a distance-based pruning step to remove grasp pairs that are too close to each other. Following this filtering process, we are left with an average of 30,000 to 80,000 dual-arm grasp pairs per object. These candidate grasp pairs are then passed through our Optimizer-based Force Closure framework to ensure that only physically stable and feasible grasps are retained.

### B. Optimization-based Force Closure Formulation

The quality of the generated set of candidate grasp pairs is then evaluated by solving the force closure optimization problem. The goal of this optimization is to determine the contact forces that can resist an external wrench ( $\mathbf{w}_{\text{ext}}$ ) acting on the object at the centre of mass in the object frame, while

satisfying the friction cone constraints and thus obeying force closure.

The dual-arm setup equipped with two dedicated two-fingered rigid grippers, results in four contact points on the object’s surface while grasping. Let the position and orientation of the contact frame relative to the object frame be represented as  $\mathbf{p}_{l/r,i} \in \mathbb{R}^3$  and  $\mathbf{R}_{l/r,i} \in \mathbb{SO}(3)$ , where  $l/r$  represents the left/right grasp and  $i$  denotes the  $i^{\text{th}}$  contact point of the respective gripper, with  $i = \{1, 2\}$ .

The force applied by the  $i^{\text{th}}$  contact point on the object in the contact frame is given by  $\mathbf{f}_{l/r,i} = \mathbf{f}_{t_{l/r,i}} + \mathbf{f}_{n_{l/r,i}}$ , where  $\mathbf{f}_{t_{l/r,i}}$  and  $\mathbf{f}_{n_{l/r,i}}$  represent the tangential and normal components of the contact forces, respectively. The grasp matrix  $\mathbf{G}_{l/r,i}$  maps the contact force  $\mathbf{f}_{l/r,i}$  to the object frame and is given by

$$\mathbf{G}_{l/r,i} = \begin{bmatrix} \mathbf{R}_{l/r,i} & \mathbf{0} \\ [\mathbf{p}_{l/r,i}]_{\times} \mathbf{R}_{l/r,i} & \mathbf{R}_{l/r,i} \end{bmatrix} \cdot \mathbf{B},$$

where  $\mathbf{B} = [\mathbf{I}_{3 \times 3} \ \mathbf{0}_{3 \times 3}]^T$  represents the contact basis for

the Point Contact With Friction (PCWF) model [22] and  $[\mathbf{p}_{l/r,i}]_{\times} \in \mathbb{R}^{3 \times 3}$  is the skew-symmetric matrix of the contact position  $\mathbf{p}_{l/r,i}$ .

Candidate grasp pairs are validated using a rank test to ensure grasp feasibility. The grasp matrix  $\mathbf{G}$  is formed by concatenating individual contact grasp matrices. Grasp pairs whose grasp matrix  $\mathbf{G}$  is not full rank are considered infeasible, indicating insufficient wrench controllability.

The remaining grasp pairs proceed to the next stage, where the optimal force distribution is determined by formulating the force closure optimization problem as a Second-Order Cone Program (SOCP). The objective function is:

The objective function is:

$$\mathcal{L} = \left\| \left( \sum_{i=1}^2 \mathbf{G}_{l/r,i} \cdot \mathbf{f}_{l/r,i} \right) + \mathbf{w}_{\text{ext}} \right\|_2^2 \quad (1)$$

The optimization problem is formulated as:

$$\min_{\mathbf{f}_{l,1}, \mathbf{f}_{l,2}, \mathbf{f}_{r,1}, \mathbf{f}_{r,2}} \mathcal{L}, \quad (2)$$

subject to,

$$\|\mathbf{f}_{l/r,i}\|_2 \leq \mathbf{f}_{\text{high}}, \quad \forall i \in \{1, 2\} \quad (3)$$

$$\|\mathbf{f}_{t_{l/r,i}}\|_2 \leq \mu \mathbf{f}_{n_{l/r,i}}, \quad \forall i \in \{1, 2\} \quad (4)$$

Constraint (3) ensures that the contact forces are within the physical limits of the grippers and  $\mathbf{f}_{\text{high}}$  is the maximum contact force the gripper can exert.

### C. Dual-arm Grasps Selection

The optimization problem is solved using a convex optimization solver, CVXPY [36]. The quality of the grasp is evaluated based on the optimal value of the objective function, which indicates how well the grasp can resist the external wrench. We define a grasp to be successful if the optimization loss satisfies,

$$\mathcal{L} \leq 10^{-5} \quad (5)$$

This threshold ensures that the net wrench on the object is close to zero, indicating a stable grasp configuration. Grasps failing to meet this criterion are discarded to maintain dataset quality. Finally, to construct a reliable and diverse dataset, we retain successful grasp pairs that satisfy both the rank condition and the force closure condition per object. This selection process ensures that only high-quality, physically feasible grasps are included, making the dataset well-suited for benchmarking dual-arm grasping algorithms.

The force closure formulation in the existing dataset differs from our approach in several key aspects. DA2 does not compute optimal force values explicitly, focusing instead on generating grasp candidates through antipodal sampling and evaluating their stability through force closure. In contrast, our formulation directly incorporates the computation of optimal contact forces, ensuring that the grasp can resist external disturbances while satisfying friction cone constraints.

Furthermore, our approach explicitly accounts for the physical limits of gripper forces, making it more practical for real-world applications where gripper capabilities play a critical role. By incorporating these constraints, we ensure that the generated grasps are not only theoretically stable but also physically feasible, allowing for reliable execution within the hardware’s operational limits. The superior performance of classifiers trained on our dataset in physics simulations further validates the effectiveness of our dataset in enabling robust and stable dual-arm grasps.

## IV. BENCHMARK DATASET CREATION

Benchmarking grasping performance is challenging due to variations in controllers, grippers, and robotic arms. While it is practically infeasible to be agnostic to gripper designs, we ensure that a successful grasp renders the object immovable under external forces. To establish this benchmark, we define the essential criteria for grasp stability by ensuring that the grasp prevents object displacement when subjected to external disturbances. Notably, grasps that appear successful based on contact points alone do not always translate into stable grasps, as they often involve area contact rather than point contact.

To construct our benchmark dataset, we select 300 objects from our object list, designating them as part of the unseen object split. This subset serves as the test set for subsequent experiments, allowing us to evaluate grasp stability.

We employ the grasp generation technique outlined in Section III to generate valid dual-arm grasps for these objects. These grasps are then evaluated in the Issac Gym simulator [15] using physics-based validation. During evaluation, the Franka Panda grippers are initialized at the designated grasp pose with gravity disabled to prevent premature object displacement. Each gripper can exert a maximum gripping force of 70N<sup>1</sup>. Once the grippers fully close, gravity is enabled, and forces are applied to stabilize the object. A grasp is deemed successful if the object remains stationary under these conditions. Otherwise, it is classified as a failure.

This evaluation process establishes a ground truth dataset, independent of predefined grasp quality metrics or assumptions, providing an objective and practical measure of grasp success. As a result, we compile a benchmark dataset of 30,000 dual-arm grasps, including their associated objects and simulation-verified ground truth labels.

## V. EXPERIMENTS AND RESULTS

In this section, we present the evaluation metrics for dual-arm grasps and analyze the performance of our proposed baselines on the DG16M and DA2 datasets.

**Metrics:** We evaluate grasps using two complementary approaches: **Force Closure Evaluation (FCE)** and **Grasp Success Rate (GSR)**. These metrics measure the feasibility and robustness of grasps under static and dynamic conditions, respectively.

*Force Closure Evaluation (FCE):* This evaluation is an analytical method that assesses grasp feasibility through

<sup>1</sup>Franka Emika Robot’s Instruction Handbook

Model	Our Dataset		DA2 Dataset	
	FCE(%) $\uparrow$	GSR(%) $\uparrow$	FCE(%) $\uparrow$	GSR(%) $\uparrow$
Dual-PointNetGPD	69.5	61.22	61.8	53.26
<b>CGDF-Classifier</b>	<b>88.73</b>	<b>76.34</b>	<b>65.25</b>	<b>54.33</b>

TABLE II: Performance comparison of different models on our dataset and DA2 dataset.

force closure analysis. The optimizer computes the required contact forces at the grasp points, ensuring static equilibrium while satisfying friction constraints. A grasp is considered successful if these forces can counteract external disturbances and maintain the object’s stability.

*Grasp Success Rate (GSR)*: This evaluation method simulates real-world grasp execution using floating grippers in Isaac Gym Simulator [15], as described in Section IV. It measures not only the theoretical feasibility of grasps but also their practical stability, ensuring they are viable for real-world execution.

**Baseline methods**: To assess the efficacy of our dual-arm grasping framework, we benchmark it against two baseline grasp classification methods: (1) We adapt PointNetGPD [34] to a dual-arm setting as done in [14]. (2) We use SE(3)-aware grasp diffusion models [33], [37] to give us grasp descriptors and train a classifier head on the features. We call these two baselines Dual-PointNetGPD, and CGDF-Classifier.

We use the Dual-PointNetGPD architecture, as described in [14]. To adapt the original PointNetGPD architecture for evaluating dual-arm grasps, firstly, the gripper templates are transformed to their grasp poses using transformation matrix  $\mathbf{H} \in SE(3)$ . To obtain grasp region information, the nearest 512 points from the object point cloud  $P_0$  are selected around each transformed gripper pose, forming two local point cloud regions,  $P_1, P_2$ . The combined set of sampled points,  $P = P_1 \cup P_2$ , serves as a local grasp representation, which is then passed through the PointNetGPD architecture to extract a feature vector. This feature vector is subsequently processed by a Multi-Layer Perceptron (MLP) to classify whether the grasp is successful or not.

Recent approaches have explored incorporating both object geometry and grasp descriptors into the learning process. For the CGDF-Classifier, we build upon [33] by integrating a classification head following the vision encoder and feature encoder outputs, which is based on neural descriptor fields [38]. The network receives the object point cloud  $P_o \in \mathbb{R}^{M \times 3}$ , consisting of  $M$  points, and the gripper template point cloud  $P_g \in \mathbb{R}^{N \times 3}$ , consisting of  $N$  points. The gripper templates are transformed into the object’s plane using the transformation matrix  $\mathbf{H} \in SE(3)$ . This design enables the model to learn both object features and the relative grasp information within the object’s reference frame. Based on these feature vectors, the classifier determines whether a given grasp is successful or not.

**Results**: As shown in Table II, models trained on our

Generation Technique	Evaluation Metrics	
	FCE(%) $\uparrow$	GSR(%) $\uparrow$
Random Dual-Arm Pairs	16.33	23.66
Farthest-Grasp Pairs	26.16	38.45
<b>DG16M (Ours)</b>	<b>100.0</b>	<b>76.33</b>

TABLE III: Ablation study comparing different dual-arm grasp generation techniques.

dataset consistently outperform those trained on DA2 across both the metrics. The higher performance of both the models on our dataset highlights the superior quality of our dataset, which employs well-defined force closure formulation. This results in well separated positive and negative grasp pairs, enabling models to learn more effectively.

A key observation is that models trained on DA2 exhibit near-random classification performance in simulation, as indicated by GSR values hovering around 50%. This suggests that DA2-trained models struggle to reliably distinguish between successful and failed grasps, classifying them almost at random. Although DA2 incorporates a structured grasp stability criterion, its grasp distributions may not transfer effectively to simulation-based evaluations, potentially limiting generalization. In contrast, our dataset provides a better grasp distribution, ensuring that models learn from more stable and physically meaningful examples. Unlike DA2-trained models, our models not only achieve high theoretical grasp feasibility (FCE) but also demonstrate strong performance in realistic physics simulations (GSR), proving their viability beyond just theoretical metrics.

Additionally, we observe a slight performance gap between FCE and GSR, largely due to the challenges of translating theoretical grasp stability into practical execution in a simulated environment. Factors such as unmodeled physical effects, contact dynamics, and complex gripper-object interactions contribute to this discrepancy.

Further, we validate the quality of our dataset by analyzing grasp stability,  $Q(\mathbf{G}) = \sqrt{\det(\mathbf{G}\mathbf{G}^T)}$ , where  $\mathbf{G}$  is the grasp matrix, which quantifies the volume of the grasp wrench space, reflecting the grasp’s ability to transfer contact forces into wrenches that stabilize the object [39]. A large  $Q(\mathbf{G})$  indicates a robust grasp with good manipulability, while a near-zero  $Q(\mathbf{G})$  suggests a failing grasp that is close to a singular configuration and unable to resist external disturbances. As shown in Figure 4, our dataset (purple bars) exhibits a clear separation between successful and failed grasps, with most failures and very few successes having near-zero determinants. In contrast, the DA2 dataset (pink bars) shows little distinction between passed and failed grasps, as its determinant values do not follow the expected trend. Many successful and failed grasps alike exhibit near-zero determinants, undermining their reliability as grasp stability indicators.

**Qualitative Results**. Figure 3 provides a qualitative comparison of grasps generated by models trained on our dataset

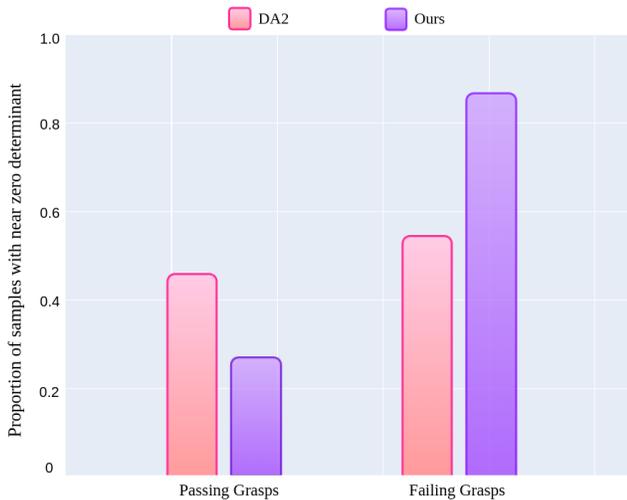


Fig. 4: **Grasp Stability Analysis.** We compute grasp stability using  $Q(\mathbf{G}) = \sqrt{\det(\mathbf{G}\mathbf{G}^T)}$ , which measures the volume of the grasp wrench space. Our dataset follows the expected trend of having large  $Q(\mathbf{G})$  values for passing grasps while unstable grasps having near-zero  $Q(\mathbf{G})$  values. However, DA2 shows both stable and unstable grasps with similar low  $Q(\mathbf{G})$  values, indicating less manipulable grasps.

versus DA2. For large objects with complex geometries, models trained on DA2 often detect grasps that fail to satisfy force closure or dynamic stability conditions. In contrast models trained on our dataset detect grasps that are both theoretically feasible and practically executable. DA2-trained models often produce unsuccessful grasps because the dataset’s grasp samples frequently fail to secure the object or maintain contact. This instability arises from factors such as grasp pairs being too close together, positioned on the same side of the center of mass, or lacking proper force balance, all of which compromise grasp stability. These failures are primarily due to the relaxed force closure condition used in DA2, which does not account for optimal contact forces and physical limits of the grippers, leading to infeasible grasps.

**Ablations.** Our key design decision in generating dual-arm grasps is the use of an optimizer-based force closure formulation to select stable grasp pairs from candidate samples. To evaluate the necessity and effectiveness of this approach, we conduct ablation studies comparing our grasp selection method against two alternative strategies. (1) Random Pairing of Single-Arm Grasps, where single-arm grasps are sampled and randomly combined to form paired grasps, and (2) Diametrically Opposite Pairing, where single-arm grasps are sampled, and the farthest grasps are paired together. The results are presented in Table III.

- 1) **Random Pairing of Single-Arm Grasps:** After sampling grasps through antipodal sampling, random pairing is performed. This naive approach results in a low FCE of 16.33% and GSR of 23.66%, indicating that most grasp pairs fail to maintain stability. Since the pairing is done arbitrarily, the resulting grasps often lack proper force balance and fail to satisfy stability constraints, leading to frequent failures in

both theoretical evaluation and simulation tests. This ablation highlights why merely relying on antipodal sampling is insufficient and demonstrates the necessity of incorporating force closure to ensure stable dual-arm grasps.

- 2) **Pairing Diametrically Opposite Grasps:** Ensuring that grasp pairs are geometrically opposite improves FCE to 26.16% and GSR to 38.45%. While this method guarantees spatial separation, it incorrectly assumes that geometric opposition alone ensures stability. In reality, it ignores force and torque equilibrium, causing many grasp pairs to still fail under external disturbances.
- 3) **Force Closure-Based Pairing:** Our method, which explicitly enforces force closure constraints, achieves 100% FCE and a GSR of 76.33%, significantly outperforming both alternative approaches. By ensuring that grasp pairs satisfy both spatial separation and physical stability, our approach produces grasps that are theoretically stable and also successful in simulated environments.

The ablation studies validate our force closure approach. Random pairing lacks physical reasoning, while diametrically opposite pairing improves but remains limited. Force closure, by incorporating physical constraints, delivers the best performance, underscoring its importance in dual-arm grasp generation.

## VI. CONCLUSIONS AND FUTURE WORK

We introduced DG16M, a large-scale dual-arm grasp dataset with 16 million grasps, addressing limitations in existing datasets like DA2. DG16M enforces force-closure constraints and incorporates simulation-based validation to ensure both theoretical and physical grasp stability. Additionally, we provide a benchmark of 30,000 simulation-verified grasps across 300 objects for robust evaluation.

Deep learning-based grasp classifiers trained on DG16M, including Dual-PointNetGPD and CGDF-Classifier, achieved significantly higher Grasp Success Rates (GSR) and Force Closure Evaluation (FCE) scores than those trained on DA2. This underscores the importance of structured, high-quality grasp distributions in improving generalization across diverse objects and configurations, making DG16M a strong foundation for dual-arm grasping research.

Future work could explore real-world grasp trials with DG16M to validate its effectiveness beyond simulation. Additionally, it could investigate adaptive grasping strategies that dynamically adjust parameters based on object properties and external disturbances, enhancing the versatility and reliability of dual-arm robotic systems in real-world tasks.

## REFERENCES

- [1] J. Grannen, Y. Wu, B. Vu, and D. Sadigh, “Stabilize to act: Learning to coordinate for bimanual manipulation,” in *Conference on Robot Learning*. PMLR, 2023, pp. 563–576.
- [2] A. Billard and D. Kragic, “Trends and challenges in robot manipulation,” *Science*, vol. 364, p. eaat8414, 06 2019.
- [3] A. Edsinger and C. C. Kemp, “Two arms are better than one: A behavior based control system for assistive bimanual manipulation,” 2008.

- [4] M. Abbas, J. Narayan, and S. K. Dwivedy, "A systematic review on cooperative dual-arm manipulators: modeling, planning, control, and vision strategies," *International Journal of Intelligent Robotics and Applications*, 2023.
- [5] C. Smith, Y. Karayiannidis, L. Nalpantidis, X. Gratal, P. Qi, D. V. Dimarogonas, and D. Kragic, "Dual arm manipulation—a survey," *Robotics and Autonomous Systems*, vol. 60, no. 10, pp. 1340–1353, 2012.
- [6] M. Sundermeyer, A. Mousavian, R. Triebel, and D. Fox, "Contact-graspnet: Efficient 6-dof grasp generation in cluttered scenes," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 13 438–13 444.
- [7] A. Mousavian, C. Eppner, and D. Fox, "6-dof graspnet: Variational grasp generation for object manipulation," 2019.
- [8] J. Lundell, F. Verdoja, T. N. Le, A. Mousavian, D. Fox, and V. Kyrki, "Constrained generative sampling of 6-dof grasps," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023, pp. 2940–2946.
- [9] 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," *Trans. Rob.*, vol. 39, no. 5, p. 3929–3945, Oct. 2023. [Online]. Available: <https://doi.org/10.1109/TRO.2023.3281153>
- [10] A. Vuong, M. N. Vu, H. Le, B. Huang, B. P. K. Huynh, T. D. Vo, A. Kugi, and A. Nguyen, "Grasp-anything: Large-scale grasp dataset from foundation models," *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 14 030–14 037, 2023. [Online]. Available: <https://api.semanticscholar.org/CorpusID:262045996>
- [11] J. Mahler, J. Liang, S. Niyaz, M. Laskey, R. Doan, X. Liu, J. A. Ojea, and K. Goldberg, "Dex-net 2.0: Deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics," 2017.
- [12] A. Miller and P. Allen, "Graspit! a versatile simulator for robotic grasping," *IEEE Robotics Automation Magazine*, vol. 11, no. 4, pp. 110–122, 2004.
- [13] C. Eppner, A. Mousavian, and D. Fox, "Acronym: A large-scale grasp dataset based on simulation," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021.
- [14] G. Zhai, Y. Zheng, Z. Xu, X. Kong, Y. Liu, B. Busam, Y. Ren, N. Navab, and Z. Zhang, "Da<sup>2</sup> dataset: Toward dexterity-aware dual-arm grasping," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 8941–8948, 2022.
- [15] V. Makoviychuk, L. Wawrzyniak, Y. Guo, M. Lu, K. Storey, M. Macklin, D. Hoeller, N. Rudin, A. Allshire, A. Handa *et al.*, "Isaac gym: High performance gpu-based physics simulation for robot learning," *arXiv preprint arXiv:2108.10470*, 2021.
- [16] R. Diankov, "Automated construction of robotic manipulation programs," 2018.
- [17] D. Kappler, J. Bohg, and S. Schaal, "Leveraging big data for grasp planning," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 4304–4311.
- [18] A. Miller, S. Knoop, H. Christensen, and P. Allen, "Automatic grasp planning using shape primitives," in *2003 IEEE International Conference on Robotics and Automation*, vol. 2, 2003, pp. 1824–1829 vol.2.
- [19] M. Veres, M. Moussa, and G. Taylor, "An integrated simulator and dataset that combines grasping and vision for deep learning," 02 2017.
- [20] C. Eppner, A. Mousavian, and D. Fox, "A billion ways to grasp: An evaluation of grasp sampling schemes on a dense, physics-based grasp data set," in *The International Symposium of Robotics Research*. Springer, 2019, pp. 890–905.
- [21] A. ten Pas and R. Platt, "Using geometry to detect grasps in 3d point clouds," 2015. [Online]. Available: <https://arxiv.org/abs/1501.03100>
- [22] R. M. Murray, S. Sastry, and Z. Li, "A mathematical introduction to robotic manipulation," in *Proceedings of the Conference on Robotics Manipulation*, 1994. [Online]. Available: <https://api.semanticscholar.org/CorpusID:108605633>
- [23] V.-D. Nguyen, "Constructing force-closure grasp," *Int. J. Robotics Res.*, vol. 7, no. 3, pp. 157–171, 1988.
- [24] 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.
- [25] D. Guo, Y. Xiang, S. Zhao, X. Zhu, M. Tomizuka, M. Ding, and W. Zhan, "Phygrasp: generalizing robotic grasping with physics-informed large multimodal models," *arXiv preprint arXiv:2402.16836*, 2024.
- [26] 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.
- [27] Y. Jiang, S. Moseson, and A. Saxena, "Efficient grasping from rgb-d images: Learning using a new rectangle representation," 05 2011, pp. 3304–3311.
- [28] A. Depierre, E. Dellandréa, and L. Chen, "Jacquard: A large scale dataset for robotic grasp detection," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 3511–3516.
- [29] H. Zhang, D. Yang, H. Wang, B. Zhao, X. Lan, J. Ding, and N. Zheng, "Regrad: A large-scale relational grasp dataset for safe and object-specific robotic grasping in clutter," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 2929–2936, 2022.
- [30] B. Calli, A. Singh, A. Walsman, S. Srinivasa, P. Abbeel, and A. M. Dollar, "The ycb object and model set: Towards common benchmarks for manipulation research," in *Proceedings of the 2015 IEEE International Conference on Advanced Robotics (ICAR)*, 2015.
- [31] L. Downs, A. Francis, N. Koenig, B. Kinman, R. M. Hickman, K. Reymann, T. B. McHugh, and V. Vanhoucke, "Google scanned objects: A high-quality dataset of 3d scanned household items," *2022 International Conference on Robotics and Automation (ICRA)*, pp. 2553–2560, 2022. [Online]. Available: <https://api.semanticscholar.org/CorpusID:248392390>
- [32] A. X. Chang, T. Funkhouser, L. Guibas, P. Hanrahan, Q. Huang, Z. Li, S. Savarese, M. Savva, S. Song, H. Su *et al.*, "Shapenet: An information-rich 3d model repository," *arXiv preprint arXiv:1512.03012*, 2015.
- [33] G. Singh, S. Kalwar, M. F. Karim, B. Sen, N. Govindan, S. Sridhar, and K. M. Krishna, "Constrained 6-dof grasp generation on complex shapes for improved dual-arm manipulation," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2024, pp. 7344–7350.
- [34] H. Liang, X. Ma, S. Li, M. Görner, S. Tang, B. Fang, F. Sun, and J. Zhang, "PointNetGPD: Detecting grasp configurations from point sets," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2019.
- [35] M. F. Karim, S. Bollimuntha, M. S. Hashmi, A. Das, G. Singh, S. Sridhar, A. K. Singh, N. Govindan, and K. M. Krishna, "Davi: Adaptive dual-arm manipulation with reinforcement learning and variable impedance control," *arXiv preprint arXiv:2410.19712*, 2024.
- [36] S. Diamond and S. Boyd, "Cvxpy: A python-embedded modeling language for convex optimization," *Journal of Machine Learning Research*, vol. 17, no. 83, pp. 1–5, 2016.
- [37] J. Urain, N. Funk, J. Peters, and G. Chalvatzaki, "Se(3)-diffusionfields: Learning smooth cost functions for joint grasp and motion optimization through diffusion," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 5923–5930.
- [38] A. Simeonov, Y. Du, A. Tagliasacchi, J. B. Tenenbaum, A. Rodriguez, P. Agrawal, and V. Sitzmann, "Neural descriptor fields: SE(3)-equivariant object representations for manipulation," in *Proceedings of the 2022 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 6394–6400.
- [39] P. Song, J. A. C. Ramón, and Y. Mezouar, "Dynamic evaluation of deformable object grasping," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4392–4399, 2022.