

A Billion Ways to Grasp: An Evaluation of Grasp Sampling Schemes on a Dense, Physics-based Grasp Data Set

Clemens Eppner, Arsalan Mousavian, and Dieter Fox

NVIDIA USA,
ceppner@nvidia.com

Abstract. Robot grasping is often formulated as a learning problem. With the increasing speed and quality of physics simulations, generating large-scale grasping data sets that feed learning algorithms is becoming more and more popular. An often overlooked question is how to generate the grasps that make up these data sets. In this paper, we review, classify, and compare different grasp sampling strategies. Our evaluation is based on a fine-grained discretization of SE(3) and uses physics-based simulation to evaluate the quality and robustness of the corresponding parallel-jaw grasps. Specifically, we consider more than 1 billion grasps for each of the 21 objects from the YCB data set. This dense data set lets us evaluate existing sampling schemes w.r.t. their bias and efficiency. Our experiments show that some popular sampling schemes contain significant bias and do not cover all possible ways an object can be grasped. The data is available at <https://sites.google.com/view/abillionwaystograsp/>.

Keywords: robotic grasping, manipulation, simulation, sampling

1 Introduction

Grasping is a fundamental skill for any robotic manipulation system. Most commonly it is solved in a data-driven fashion (Bohg et al, 2013), either through supervision (Mahler et al, 2017) or reinforcement (Levine et al, 2016; Kalashnikov et al, 2018). To satisfy the data hunger of these learning methods, grasps are often labeled in simulation.

The advantages of evaluating grasps in simulation are manifold: data collection can be scaled easily, grasp conditions can be controlled, robots won't break, resetting is trivial, and the supervision signal benefits from a fully observable environment. Although the gap between simulation and reality needs to be addressed, it has been shown that models trained exclusively with synthetic grasp data can perform successfully in the real world (Mahler et al, 2017).

Generating synthetic grasp data is usually based on heuristics that select a gripper pose relative to the object. Oftentimes these heuristics don't get much attention in grasp learning publications and occupy only a minor paragraph. In this paper, we thoroughly analyze and compare these different sampling schemes.

We focus our evaluation on quantifying the grasp coverage for each heuristic. Do some sampling schemes cover the space of all possible grasps of an object better than others?

To answer this question empirically we need a ground truth grasp density for a given object. We acquire this ground truth by discretizing the space of all possible grasps with high resolution and evaluate them in a physics simulation (Macklin et al, 2014). This reference data set contains dense grasp sets for 21 objects of the YCB object set (Calli et al, 2017). To the best of our knowledge there has been no prior attempt to exhaustively describe all possible grasps for an object.

Note, that our analysis is not limited by the realism of the results produced in the physics simulation. However, we show that the simulated grasps can be successfully executed on a real robotic system (Sec. 5.2). Furthermore, due to the denseness of the data we can generate robust versions of the original grasp sets and use those for evaluation.

Contribution Our contribution is two-fold:

1. We present a set of all possible parallel-jaw grasps for 21 objects. It is generated by discretizing $\text{SE}(3)$ with a resolution of (5 mm, 7.5°) and executing more than a billion grasps in a physics simulator.
2. We use this data set to study and compare different sampling schemes for grasping. Our comparison shows how grasp coverage is affected by using different samplers. Based on these results we recommend sampling strategies for generating large-scale grasp data sets.

Organization The paper is structured as follows. First, we review existing grasp sampling schemes that are used for generating data sets. We then categorize those methods into a coherent taxonomy, and present our evaluation criteria. Finally, we compare a number of sampling schemes and discuss their pros and cons.

2 Related Work

To the best of our knowledge, there has been no prior evaluation of different sampling strategies for generating grasp data sets. In contrast, the field of sampling-based motion planning (LaValle, 2006) is rife with sampling strategies based on heuristics. However, it has been shown that no sampling strategy outperforms all others in all scenarios (Lindemann and LaValle, 2005; Elbanhawi and Simic, 2014).

In the following we briefly review grasp sampling strategies, clustered according to their use case. Since our evaluation uses a physics simulator, we also give a brief overview about the usage of physics simulation for robot grasping.

Grasp Sampling for Planning/Inference Sampling-based techniques are often used to find optimal grasps given an analytical or learned model of grasp

quality. Examples are simulated annealing (Ciocarlie et al, 2007; Hang et al, 2016) or the cross-entropy method (Mahler et al, 2017; Yan et al, 2017). We do not include these black-box optimization methods in our analysis for two reasons. First, these methods are used for finding an optimum, while we are interested in discovering the entire grasp distribution. Second, since these sampling approaches require models that can be evaluated quickly, their performance also depends on the quality of the model approximation.

Grasp Sampling for Generating Real-World Data Sets The Cornell grasping dataset Jiang et al (2011) contains $\approx 5K$ human-labelled grasps for 280 objects which are represented as rectangles in the image plane. Since the data is relatively sparse (≈ 18 grasps per object), human-sampled, and viewpoints correlate with the object’s equilibrium poses, bias is inevitable. Pinto and Gupta (2016) collect $50K$ top-down grasps autonomously by sampling random grasp points and orientations in $SE(2)$. Again, the sampling method is biased since sampling only happens in planes parallel to the object’s equilibrium poses.

Grasp Sampling for Generating Synthetic Datasets Our main focus is on sampling methods that are used to generate large-scale synthetic datasets via physics simulation. The Columbia grasp database (Goldfeder et al, 2008) generated grasps using the simulated annealing-based Eigengrasp planner (Ciocarlie et al, 2007) described above.

Zhou and Hauser (2017) learn from grasping data generated in a physics simulation. Their sampling scheme samples random lines that intersect the object’s center of mass. The hand pose is then determined by shifting along this line in an arbitrary orientation. This scheme ensures that the space near the object’s COM is more densely sampled than poses that are further away.

Kappler et al (2015) present a large dataset of grasps generated by the sampling scheme of Diankov (2010). It samples hand approach vectors close to the object’s surface normals and chooses random roll angles and standoff distances. In a similar vein, the grasp dataset by Kleinhans et al (2015) uses surface normals to sample the hand position while the orientation is chosen randomly. Veres et al (2017) sample grasp poses around the object such that the hand’s approach vector intersects the object’s bounding box. In the following section (Sec. 3) we will group all these sampling methods as approach-based schemes.

In contrast to approach-based samplers are antipodal-based sampling schemes. They sample directly the potential contact points with the object. Examples are the data sets used in (Mahler et al, 2017; ten Pas and Platt, 2018; Liang et al, 2018).

Physics Simulation for Grasping Our analysis relies on the evaluation of grasp candidates in a physics simulation (Macklin et al, 2014). Simulating stable grasps is challenging (Erez et al, 2015) and the results never fully transfer to the real world (Collins et al, 2018). However, it has been shown that simulation provides significantly more information about grasp success than traditional grasp

Grasp Result	Guided by Object Geometry	Category	Publications generating Grasp Data Sets
\times	\times	Uniform	
		Non-uniform	Zhou and Hauser (2017)
	\checkmark	Approach-based	Kappler et al (2015); Kleinhans et al (2015); Veres et al (2017)
		Antipodal-based	Mahler et al (2017); ten Pas and Platt (2018)
\checkmark	\checkmark	Adaptive	Goldfeder et al (2008)

Table 1. A taxonomy of grasp samplers. See text for detailed explanation.

quality metrics, such as force-closure analysis Kim et al (2013). This is due to focusing on the entire grasp process including object dynamics, instead of only measuring the quality of the established contacts.

3 A Taxonomy of Sampling Strategies for Grasping

We define a grasp as a combination of a pre-grasp and a closing motion. The pre-grasp $g \in SE(3) \times \mathbb{R}^n$ describes the pose and configuration of the hand (n is the number of internal DoF) prior to the execution of a controller that represents the closing motion. In the remainder we assume that the closing controller position-controls the hand pose and uses a force-based control law to close the fingers. Given a fixed closing controller, we will focus our evaluation on parallel-jaw grippers ($n = 1$), i.e., the space of all possible grasps is $SE(3) \times \mathbb{R}$. We assume the fingers to be maximally opened during the pre-grasp which further reduces the grasp space to $SE(3)$.

We compare different sampling methods of this grasp space regarding to how well they cover all successful grasps. In the previous section we reviewed those samplers according to their application scenario, i.e. whether they were used to generate data or for planning. Now, we want to subdivide them in more detail into categories based on what information they use and how they parameterize the grasp space. The taxonomy shown in Tab. 1 classifies grasp sampling along the following criteria:

Guided by Grasp Result Our first broad distinction is whether a grasp sampler evaluates the grasp quality function and uses this outcome when drawing subsequent samples. This is independent of the actual realization of the grasp quality function. It could be any classical grasp metric (Roa and Suárez, 2015), a physics simulation, or even the physical execution of the grasp on a real platform.

We focus our empirical evaluation on sampling methods that are not guided by this information. Since most grasp quality functions depend on contact and are noncontinuous, grasp information is very local which is of limited value when generating datasets that contain diverse grasps that should fully cover an object.

The large majority of existing grasp datasets is generated by this methods that are not guided by the grasp outcome.

Guided by Object Geometry Most grasp sampling methods are guided by surface information of the object. This is often done by parameterizing the grasp using surface normals, either by aligning the hand’s approach vector (or the palm’s surface normal) with the object’s surface normal or by aligning the expected finger contact normals with object’s surface. We will show in our empirical analysis that although these methods are effective at generating grasps, they are biased, i.e., the resulting grasps do not fully cover all possible grasps of an object.

Uniform Samplers Without using any geometric information about the object or the outcome of a grasp sample the best thing we can do is sampling the bounded $SE(3)$ space uniformly. The uniformity of samples can be expressed by measures like discrepancy and dispersion (LaValle, 2006). As a result multiple sequences have been proposed that result in better uniformity than those produced by pseudo-random number generators. Among low-discrepancy sampling there are three categories: Halton (Halton, 1960)/Hammersley sequences, (t,s)-sequences and (t,m,s)-nets, and lattices such as the Sukharev grid (Sukharev, 1971). Lattices are finite point sets which limits their applicability. But incremental grids for $SO(3)$ (Yershova et al, 2010) and $SE(3)$ (Lindemann et al, 2004) have been proposed.

Note, that low-discrepancy sampling techniques are not limited to uniform sampling schemes. All of the following sampling methods can benefit from applying low-discrepancy sampling for their parameters or subsets of them. But care needs to be taken, given that low discrepancy in parameter space not necessarily leads to low discrepancy in $SE(3)$. In our evaluation we include a uniform sampling scheme based on a pseudo-random number generator.

Non-uniform Samplers There are only few sampling methods that do not exploit information about the object’s geometry but still sample non-uniformly. One example is the approach taken by Zhou and Hauser (2017). They sample random lines that go through the origin (i.e. center-of-mass) of the object, with the directions being distributed uniformly. Evenly spaced points are chosen along a line that form the translation of the grasp. The orientation is sampled randomly. This scheme results in a higher density of grasp samples closer to the COM of the object.

Approach-based Samplers The majority of grasp data sets are generated via approach-based sampling methods. The approach vector of a gripper is the direction in which the grasp pose is approached and usually aligns with the palm’s surface normal. The sampling scheme most commonly aligns the approach vector with the surface normal of a randomly sampled point on the object. But there are a number of variants among those techniques. Points on the object surface are

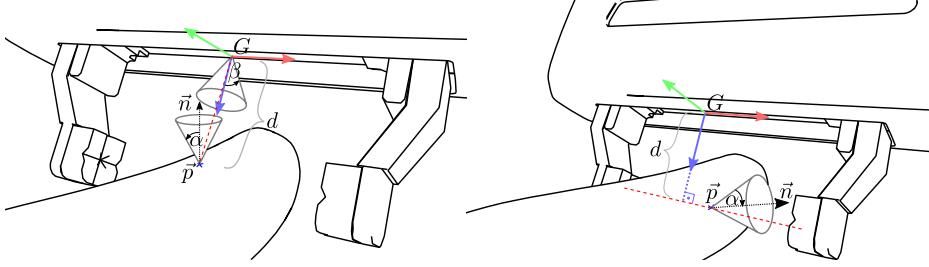


Fig. 1. Parameterization of the approach-based grasp sampling schemes (*left*) and the antipodal-based schemes (*right*). See text for details.

either sampled uniformly or selected by casting rays from a bounding box (Diankov, 2010; Kappler et al, 2015). Another approach uses the surface points and normals of a fitted primitive (box, sphere, cylinder) to sample grasps (Miller et al, 2003). Veres et al (2017) also sample the approach vector of the gripper.

For evaluation, we parameterize the most important subgroup of approach-based sampling methods as follows: Given a point on the object’s surface and its corresponding normal, a direction is chosen whose angular difference with the normal is below α , a standoff is chosen between zero and the length of the fingers, and an approach vector is chosen whose angular difference with the chosen direction is below β (see Fig. 1). The hand’s roll around the approach vector is finally chosen to be between 0 and 2π . Our evaluation contains strategies for the following (α, β) : $(0, 0)$, $(0, \pi)$, , and $(\frac{\pi}{2}, 0)$.

Antipodal-based Samplers In contrast to the approach-based heuristics, another popular group of methods tries to sample directly in the space of possible contact points between object and hand. In addition, these methods exploit the conditions under which antipodal grasps create force-closure (Mahler et al, 2017; ten Pas and Platt, 2018).

In contrast to approach-based samplers, it is non-trivial to scale antipodal-based samplers to multi-fingered hands and beyond antipodal grasps. This is due to the fact that there is no bijective mapping between hand configuration and contact locations.

For evaluation, we parameterize the antipodal-based sampling strategies as follows: Given a point on the object’s surface surface and its corresponding normal, an antipodal point is chosen by finding the farthest location of intersection with the object along a ray whose angular difference with the normal is below α . Given the two antipodal contact points, the gripper pose is defined by choosing the center point along the ray, a rotation around the ray between 0 and 2π , and a standoff in the interval $[s_{\min}, 0]$ (see Fig. 1). Our evaluation contains strategies for the following (α, s_{\min}) : $(\frac{\pi}{6}, 0)$ and $(\frac{\pi}{2}, 0)$.

Adaptive Samplers We group all strategies that select new samples based on the outcome of previous grasp samples into the category of adaptive samplers. All planning approaches described in the related work are adaptive samplers (Sec. 2). This includes methods like simulated annealing (Ciocarlie et al, 2007), cross-entropy method, importance sampling, or Bayesian optimization. We do not include any of those methods in our evaluation.

4 Evaluation

4.1 Grasp Evaluation Metrics

Our evaluation metrics are based on distances between grasps. Similar to Mahler et al (2016) we use a weighted metric. Let $g, h \in SE(3)$ be two grasps, with $g_p, h_p \in \mathbb{R}^3$ being their positions and $g_q, h_q \in \mathcal{S}^3$ their orientations represented as unit quaternions. The distance between g and h is defined as:

$$\rho(g, h) = \omega \|g_t - h_t\|_2 + \arccos(|\langle g_q, h_q \rangle|),$$

where $\omega \in \mathbb{R}$ is a weight that relates rotation and translation. Unlike Mahler et al (2016) we do not select it depending on the size of the object. Instead we keep it constant, such that a pure translation of 1 mm equals a pure rotation of 1°. Given the distance metric ρ , we now show which performance metrics we use to compare the different sampling mechanisms.

Grasp Coverage Our main objective is to find sampling methods that capture the reference grasp distribution of an object. We define different measures of grasp coverage that capture different properties as follows. The set \mathcal{X} contains all grasps sampled by a particular method, while \mathcal{R} is the reference set of all successful grasps found in simulation. Our first metric is defined as:

$$\text{cov}_1(\mathcal{X}, \mathcal{R})_\epsilon = \frac{|\{g \mid g \in \mathcal{R} \wedge \exists x \in \mathcal{X} : \rho(g, x) \leq \epsilon\}|}{|\mathcal{R}|},$$

where $\epsilon \in \mathbb{R}$ defines the maximum distance that two grasps are considered equal. Although cov_1 is intuitive it is sensitive to the choice of ϵ . It can even happen that the ordering of sampling methods according to cov_1 changes with different choices of ϵ .

To circumvent this problem we also report a grasp coverage measure based on dispersion. This metric was used in (Mahler et al, 2016) and is defined as follows:

$$\text{cov}_2(\mathcal{X}, \mathcal{R}) = \exp \left(- \max_{g \in \mathcal{R}} \min_{x \in \mathcal{X}} \rho(x, g) \right)$$

Since cov_2 is the longest of all shortest paths between \mathcal{X} and \mathcal{R} , it can be dominated by outliers in \mathcal{R} . This is possible because the reference set is generated

in a physics simulation. To get a more representative coverage measure we also report the average over all shortest paths:

$$\text{cov}_3(\mathcal{X}, \mathcal{R}) = \exp \left(-\frac{1}{|\mathcal{R}|} \sum_{g \in \mathcal{R}} \min_{x \in \mathcal{X}} \rho(x, g) \right)$$

Note, that the computational bottleneck of all coverage calculations is the nearest-neighbor search, especially since we are dealing with large sets of up to millions of elements. In our implementation we use the SE(3) k-d tree by Ichnowski and Alterovitz (2015).

Precision Oftentimes learning approaches for grasping are based on a critic or discriminative model that predicts the quality of a given grasp. Training data for such models needs to be balanced, i.e., it should roughly contain as much positive as negative grasps. Since this is not captured by the coverage metrics, we also evaluate the different sampling schemes w.r.t. their precision. Precision is defined by the ratio of successful grasps among all sampled ones.

4.2 Grasp Robustness

Grasp success is very sensitive to the accurate reproduction of the contact configuration between hand and object. Slight variations in the positioning of the hand can lead to vastly different outcomes. Grasp planning approaches have addressed this by incorporating noise models for computing grasp quality metrics (Weisz and Allen, 2012) or in physics simulations (Kim et al., 2013).

Similarly, we define the robustness of a grasp as the portion of successful grasps in its ϵ -neighborhood. Given a grasp $g \in \mathcal{R}$, a grasp set \mathcal{R} with a constant grasp density, and an indicator function $\mathbf{1}_{\text{succ}}$ denoting a successful grasp, we define:

$$\text{robust}_\epsilon(g) = \frac{|\{g \in \mathcal{R} : \rho(g, x) \leq \epsilon, \mathbf{1}_{\text{succ}}(g) = 1\}|}{|\{g \in \mathcal{R} : \rho(g, x) \leq \epsilon\}|}$$

Consequently, the robust version of a grasp set \mathcal{R} is defined as:

$$\text{robust}_{\epsilon, \gamma}(\mathcal{R}) = \left\{ g \mid g \in \mathcal{R}, \text{robust}_\epsilon(g) \geq \gamma \right\},$$

where $\gamma \in [0, 1]$ is the robustness threshold. In our evaluation we will also report the performance of the different grasp samplers w.r.t. the robust coverage metrics:

$$\text{cov}_i^{\epsilon, \gamma}(\mathcal{X}, \mathcal{R}) = \text{cov}_i(\mathcal{X}, \text{robust}_{\epsilon, \gamma}(\mathcal{R})).$$

4.3 Evaluation in Simulation

We evaluate grasps in a physics simulation. This allows us to scale our evaluation to extremely large quantities of grasp attempts (billions, in contrast to hundreds

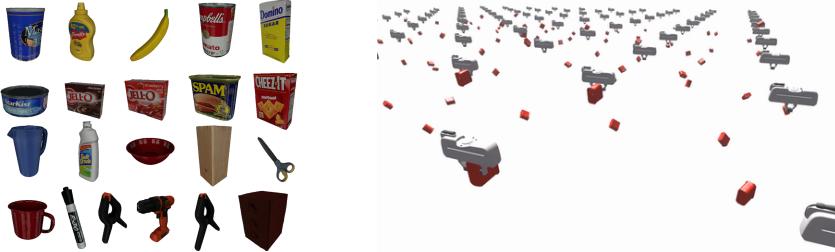


Fig. 2. Objects from the YCB dataset (left) were grasped in simulation by a parallel-jaw gripper (right).

of thousands in real-world setups (Levine et al, 2016)). It also allows us to control all aspects of the data collection process, generating dense grasp distributions for single objects. We use the physics simulator FleX (Macklin et al, 2014) and 21 object meshes of the YCB dataset (Calli et al, 2017), shown in Fig. 2. We assume a constant friction coefficient of 1.0 between the rubber pads of the Franka Panda gripper and all objects. All objects are assumed to have a constant density. The grasps are simulated in free space, without any gravity applied (similar to Zhou and Hauser (2017)). Given an initial hand position, the gripper closes its fingers (using a force-based control scheme) and executes a pre-defined motion trajectory that involves linear shaking along the approach vector and angular shaking around the finger closing direction. We record the amount of motion the object undergoes during finger closing and shaking. We also record whether the object stays between the fingers until the end of the simulation.

Note, that our analysis is not limited to the evaluation in a physics simulator. The sampling schemes could also be evaluated against a number of classical grasp quality metrics (Roa and Suárez, 2015). But given the evidence that classical metrics are very sensitive w.r.t. contact point locations and do not capture stability, we think simulation is a more realistic way to evaluate grasps. The experimental sections provides supporting evidence that the data generated in simulation is transferable to the real world.

5 Experimental Results

5.1 Physics-based Reference Data

We simulated the grasp outcomes for 21 different objects from the YCB dataset (Calli et al, 2017). Grasps are evenly spaced on a grid in $SE(3)$ with 5 mm between hand positions, and 7.5° between neighboring orientations. Evenly distributed orientations were ensured by applying the method of Yershova et al (2010). The simulation was done in FleX (Macklin et al, 2014) using a model of the 1-DOF Franka Panda gripper.

For each object, we simulated only those grasps that passed a collision test and which had a nonempty object volume between the fingers. All other grasps

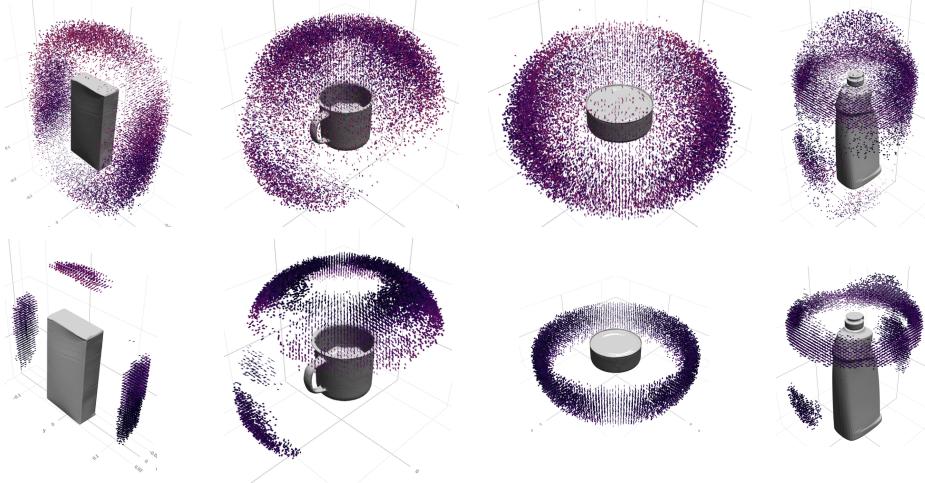


Fig. 3. Four example objects (left to right: sugar box, mug, tuna fish can, bleach cleanser) and the resulting successfully simulated grasps. Each colored point indicates a successful grasp pose. The bottom row shows robust versions of the grasp sets.

were marked as failures. In total, ≈ 317 billion grasps were sampled, of which ≈ 1 billion passed (0.32 %) the tests and were simulated in FleX. Simulations were run on 100 GPUs for one and a half months. We simulated 225 grasps in parallel on a single GPU, which lasted 90 s on average. Out of all grasps ≈ 156 million were successful (15.57 %). Fig. 3 shows successful grasps for a few objects.

5.2 Real-World Robot Experiments

To verify the simulated reference grasps, we conducted experiments in the real world with a 7-DOF Franka Panda manipulator equipped with a 1-DOF parallel-jaw gripper. Since the grasps are defined in object coordinates, we need to estimate the object poses. We use state-of-the-art object pose detectors PoseCNN (Xiang et al, 2017) and DeepIM (Li et al, 2018) to get an initial estimate and further refine it with DART (Schmidt et al, 2014) using depth.

Since it is impossible to evaluate all the reference grasps with the real robot, we verified a subset of grasps on five objects that are shown in Fig. 4. For each object, five diverse grasps are chosen and executed. Success of each grasp in these experiments depends on the accuracy of the estimated object pose, control error, and also the quality of reference grasp. For each object, 100 grasps are sampled from the robust set of grasps for each object using farthest point sampling. The grasps that lead to collisions with the support surface are removed. From the remaining grasps, five diverse grasps are chosen to be executed. Out of the 25 grasps only three failed. This is due to the discrepancy between real-world physics and simulation. For example, objects have uniform density in the simulator, they



Fig. 4. Example grasps from the simulated data set executed on the real robot.

are completely rigid and also exhibit different friction coefficients. For a video of the experiments, see <https://bit.ly/2HWEI2r>.

5.3 Comparison of Different Sampling Methods

We compared the different grasp sampling methods presented in Sec. 3. We ran each sampling method on all objects and calculated the different evaluation metrics presented in Sec. 4. We assume that a grasp pose that is in collision with the object is invalid as well as a grasp pose whose volume between the gripper’s fingers does not intersect with any part of the object. For all evaluated methods, we reject samples that do not pass these two tests.

Grasp Coverage Fig. 5 shows a comparison of all grasp samplers averaged over all objects. We show curves for two coverage metrics (cov_1 and cov_3) for the first 3 million samples and a zoom-in on the first 100,000 sampled grasps.

The uniform sampling scheme is the least biased one, attaining full coverage cov_1 within the first 3 million samples over all objects. The approach-based sampling strategies have a wide performance range depending on their parameterization. The $\text{surface}(0, 0)$ strategy is the worst, it only samples grasps along the surface normals of objects. This leads to uncovered holes in the resulting grasp set, especially close to discontinuous structures such as edges. As a result e.g. the blades of the scissors cannot be grasped from all sides equally. The $\text{surface}(\frac{\pi}{2}, 0)$ strategy does not suffer from this problem since it samples approach directions from a cone centered around the surface normals. Consequently, it is the second best sampling strategy in terms of coverage. Including the same amount of variation when choosing the gripper’s approach vector does not lead to high coverage, as shown by the curve of $\text{surface}(0, \frac{\pi}{2})$.

The antipodal-based strategies perform not as good as the best approach-based strategy. Both of them saturate at around 68 %/75 % coverage. Their bias is visualized in Fig. 6, where the reference grasps are shown that are farthest away from the ones sampled by $\text{antipodal}(\frac{\pi}{6})$. It can be seen, that grasping the lip of the meat can is not covered.

The lower plots in Fig. 5 show a magnifying view of the coverage performance during the first 100,000 samples. This is important if only a limited sampling budget is available. In this case the antipodal schemes, especially the $\text{antipodal}(\frac{\pi}{6})$ scheme is the best one. Its exploitative behavior finds suitable grasps quicker than any other sampling strategy.

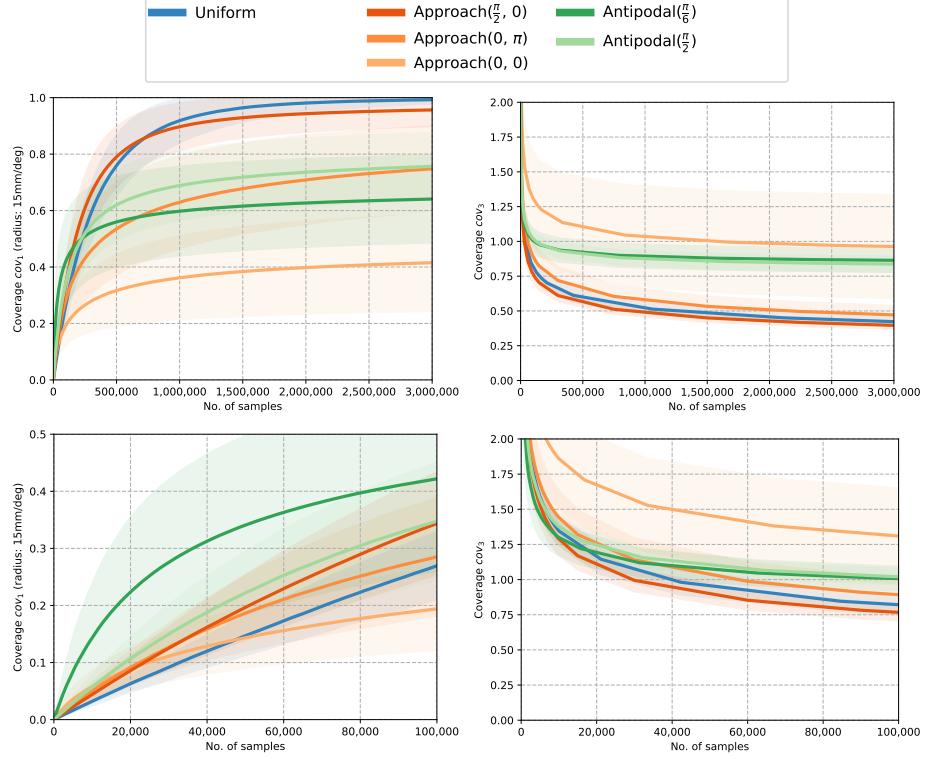


Fig. 5. Mean coverage and standard deviation over all objects for different sampling strategies. The lower plots magnify the curves during the first 100,000 samples.

Qualitative Grasp Differences The previous experiment showed that sampling heuristics that are more exploitative suffer from a high bias, i.e., they do not cover all possible grasps. But what kind of grasps are missed? To answer this question we computed the shortest distance for each successful reference grasp to the sampled grasps. Fig. 6 shows the most distant reference grasps for the antipodal($\frac{\pi}{6}$) scheme for various objects. It can be seen that the antipodal sampler misses small-scale features such as the rim of the potted meat can. It also ignores approaches directed towards edges that result in successful grasps like shown with the gelatin box, bowl, and scissor blade. See <https://bit.ly/2HWEI2r> for more examples.

Robust Grasp Coverage We evaluated the grasp samplers also w.r.t. the set of robust grasps for each object as defined in Sec. 4.2. The results shown Fig. 7 reveal that the ranking of the different heuristics does not change. Still some samplers focus more on robust grasps than others. While with fewer samples the antipodal($\frac{\pi}{6}$) scheme seems to gain the most coverage, asymptotically the approach($0, \pi$) scheme benefits the most by only considering robust grasps.

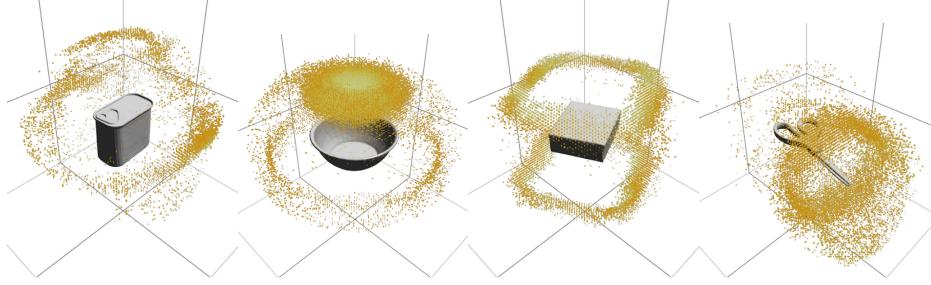


Fig. 6. Successful grasps of the potted meat can, bowl, gelatin box, and scissors that are missed by the antipodal-based sampling strategy.

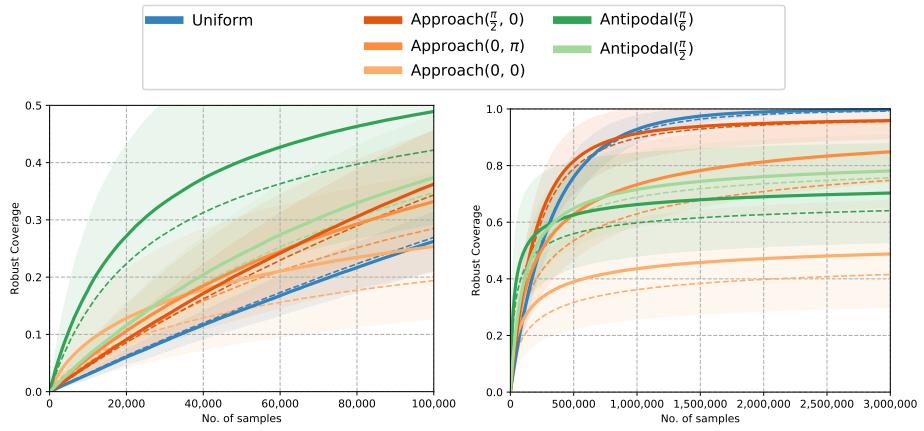


Fig. 7. Coverage for robust grasps for each sampling scheme. The dashed lines show the coverage on the original set (Fig. 5).

Precision In a final experiment we compare the precision of different sampling schemes, i.e., the probability of a sampled grasp to be successful. The results in Table 2 show that a larger bias not necessarily leads to higher precision. The uniform and approach($0, \pi$) strategy exhibit the lowest precision. The antipodal($\frac{\pi}{6}$) scheme has the highest precision. For learning approaches having a balanced set might be advantageous.

6 Discussion

Our comparison of different grasp sampling schemes exposes a kind of bias-variance dilemma. Less constrained samplers such as the uniform one will cover all successful grasps for all objects but do so at the expense of poor sample efficiency. On the other hand more constrained heuristics can be efficient but might not capture the entire subspace of possible grasps. The empirical evaluation revealed that the antipodal scheme is initially much more effective at capturing

Uniform	Approach			Antipodal	
	$(0, \pi)$	$(0, 0)$	$(\frac{\pi}{2}, 0)$	$(\frac{\pi}{6})$	$(\frac{\pi}{2})$
0.26 (± 0.15)	0.29 (± 0.17)	0.49 (± 0.28)	0.36 (± 0.23)	0.75 (± 0.26)	0.42 (± 0.24)

Table 2. Average precision (STD) of the different grasp sampling schemes.

large parts of the grasp subspace compared to the approach-based schemes. But one needs to be aware of the imposed bias. Note that for a given fixed sampling budget it is advantageous to chose a set rather than a sequence since it will lead to lower dispersion.

Our analysis focuses on the behavior of sampling heuristics as a function of the number of samples. It assumes that the computational complexity of drawing a valid sample is comparable between different heuristics. Although it is significantly more difficult, a more faithful comparison should look at the value of different heuristics per unit of computation.

Limitations Note that the simulation data has a few limitations: Due to the discretization there are aliasing effects shown by asymmetric grasp sets for symmetric objects. Additionally, we do not simulate gravity or any contact constraints with the environment. We also did not vary the internal DOF of the gripper. Adding all these dimensions would impede us from simulating all possible grasps in a reasonable amount of time.

7 Conclusions

We presented a dense data set of parallel-jaw grasps for 21 objects from the YCB data set. The data set is annotated with the results from running a physics simulation for more than a billion grasps. We showed that the quality of the simulation is reasonable, by using a model-based robotic system and transferring the successful grasps to the real world.

The data allowed us to quantify empirically for the first time the bias exposed by existing grasp sampling schemes. This will improve the understanding of data generation for 6-DOF grasp learning algorithms. Fully capturing the entire grasp distribution is important in order to plan grasps that are conditioned on task or environmental constraints and go beyond simple pick-and-place scenarios.

ACKNOWLEDGMENT We thank Miles Macklin, Viktor Makoviychuk, and Nuttapong Chentanez for support with FleX.

Bibliography

- Bohg J, Morales A, Asfour T, Kragic D (2013) Data-driven grasp synthesis survey. *IEEE Transactions on Robotics* 30(2):289–309
- Calli B, Singh A, Bruce J, Walsman A, Konolige K, Srinivasa S, Abbeel P, Dollar AM (2017) Yale-cmu-berkeley dataset for robotic manipulation research. *The International Journal of Robotics Research* 36(3):261–268
- Ciocarlie M, Goldfeder C, Allen PK (2007) Dimensionality reduction for hand-independent dexterous robotic grasping
- Collins J, Howard D, Leitner J (2018) Quantifying the reality gap in robotic manipulation tasks. arXiv preprint arXiv:181101484
- Diankov R (2010) Automated construction of robotic manipulation programs
- Elbanhawi M, Simic M (2014) Sampling-based robot motion planning: A review. *Ieee access* 2:56–77
- Erez T, Tassa Y, Todorov E (2015) Simulation tools for model-based robotics: Comparison of bullet, havok, mujoco, ode and physx. In: 2015 IEEE international conference on robotics and automation (ICRA), IEEE, pp 4397–4404
- Goldfeder C, Ciocarlie M, Dang H, Allen PK (2008) The columbia grasp database
- Halton JH (1960) On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik* 2(1):84–90
- Hang K, Li M, Stork JA, Bekiroglu Y, Pokorny FT, Billard A, Kragic D (2016) Hierarchical fingertip space: A unified framework for grasp planning and in-hand grasp adaptation. *IEEE Transactions on robotics* 32(4):960–972
- Ichnowski J, Alterovitz R (2015) Fast nearest neighbor search in for sampling-based motion planning. In: Algorithmic Foundations of Robotics XI, Springer, pp 197–214
- Jiang Y, Moseson S, Saxena A (2011) Efficient grasping from rgbd images: Learning using a new rectangle representation. In: 2011 IEEE International Conference on Robotics and Automation, IEEE, pp 3304–3311
- Kalashnikov D, Irpan A, Pastor P, Ibarz J, Herzog A, Jang E, Quillen D, Holly E, Kalakrishnan M, Vanhoucke V, et al (2018) Qt-opt: Scalable deep reinforcement learning for vision-based robotic manipulation. arXiv preprint arXiv:180610293
- Kappler D, Bohg J, Schaal S (2015) Leveraging big data for grasp planning. In: Proc. 2015 IEEE Int. Conf. on Robotics and Automation (ICRA)
- Kim J, Iwamoto K, Kuffner JJ, Ota Y, Pollard NS (2013) Physically based grasp quality evaluation under pose uncertainty. *IEEE Transactions on Robotics* 29(6):1424–1439
- Kleinhans A, Rosman BS, Michalik M, Tripp B, Detry R (2015) G3db: A database of successful and failed grasps with rgbd images, point clouds, mesh models and gripper parameters
- LaValle SM (2006) Planning algorithms. Cambridge university press
- Levine S, Pastor P, Krizhevsky A, Quillen D (2016) Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. arXiv preprint arXiv:160302199

- Li Y, Wang G, Ji X, Xiang Y, Fox D (2018) Deepim: Deep iterative matching for 6d pose estimation. In: European Conference Computer Vision (ECCV)
- Liang H, Ma X, Li S, Görner M, Tang S, Fang B, Sun F, Zhang J (2018) Pointnetgpd: Detecting grasp configurations from point sets. arXiv:180906267
- Lindemann SR, LaValle SM (2005) Current issues in sampling-based motion planning. In: 11th ISRR, Springer, pp 36–54
- Lindemann SR, Yershova A, LaValle SM (2004) Incremental grid sampling strategies in robotics. In: Algorithmic Foundations of Robotics VI, Springer, pp 313–328
- Macklin M, Müller M, Chentanez N, Kim TY (2014) Unified particle physics for real-time applications. ACM Transactions on Graphics (TOG) 33(4):153
- Mahler J, Hou B, Niyaz S, Pokorny FT, Chandra R, Goldberg K (2016) Privacy-preserving grasp planning in the cloud. In: 2016 IEEE International Conference on Automation Science and Engineering (CASE), IEEE, pp 468–475
- Mahler J, Liang J, Niyaz S, Laskey M, Doan R, Liu X, Ojea JA, Goldberg K (2017) Dex-net 2.0: Deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics. In: Proc. of Robotics: Science and Systems
- Miller AT, Knoop S, Christensen HI, Allen PK (2003) Automatic grasp planning using shape primitives
- ten Pas A, Platt R (2018) Using geometry to detect grasp poses in 3d point clouds. In: Robotics Research, Springer, pp 307–324
- Pinto L, Gupta A (2016) Supersizing self-supervision: Learning to grasp from 50k tries and 700 robot hours. In: 2016 IEEE international conference on robotics and automation (ICRA), IEEE, pp 3406–3413
- Roa MA, Suárez R (2015) Grasp quality measures: review and performance. Autonomous robots 38(1):65–88
- Schmidt T, Newcombe RA, Fox D (2014) Dart: Dense articulated real-time tracking. In: Robotics: Science and Systems, vol 2
- Sukharev AG (1971) Optimal strategies of the search for an extremum. USSR Computational Mathematics and Mathematical Physics 11(4):119–137
- Veres M, Moussa M, Taylor GW (2017) An integrated simulator and dataset that combines grasping and vision for deep learning. arXiv:170202103
- Weisz J, Allen PK (2012) Pose error robust grasping from contact wrench space metrics. In: 2012 IEEE international conference on robotics and automation, IEEE, pp 557–562
- Xiang Y, Schmidt T, Narayanan V, Fox D (2017) Posecnn: A convolutional neural network for 6d object pose estimation in cluttered scenes. arXiv preprint arXiv:171100199
- Yan X, Hsu J, Khansari M, Bai Y, Pathak A, Gupta A, Davidson J, Lee H (2017) Learning 6-dof grasping interaction via deep geometry-aware 3d representations. arXiv preprint arXiv:170807303
- Yershova A, Jain S, LaValle SM, Mitchell JC (2010) Generating uniform incremental grids on $so(3)$ using the hopf fibration. The International journal of robotics research 29(7):801–812
- Zhou Y, Hauser K (2017) 6dof grasp planning by optimizing a deep learning scoring function. In: Robotics: Science and Systems (RSS) Workshop on Revisiting Contact-Turning a Problem into a Solution