To appear in the *IEEE International Conference on Robotics and Automation*, 2014.

Combinatorial Optimization for Hierarchical Contact-level Grasping

Kaiyu Hang, Johannes A. Stork, Florian T. Pokorny and Danica Kragic

***Abstract*— We address the problem of generating force-closed point contact grasps on complex surfaces and model it as a combinatorial optimization problem. Using a multilevel refinement metaheuristic, we maximize the quality of a grasp subject to a reachability constraint by recursively forming a hierarchy of increasingly coarser optimization problems. A grasp is initialized at the top of the hierarchy and then locally refined until convergence at each level. Our approach efficiently addresses the high dimensional problem of synthesizing stable point contact grasps while resulting in stable grasps from arbitrary initial configurations. Compared to a sampling-based approach, our method yields grasps with higher grasp quality. Empirical results are presented for a set of different objects. We investigate the number of levels in the hierarchy, the computational complexity, and the performance relative to a random sampling baseline approach.**

I. Introduction

The synthesis of feasible and stable grasps on an object remains an important problem in robotics which involves sensory perception and representation [1]–[3], the encoding of task constraints [4], and high dimensional configuration spaces [5]. The planning and execution of stable grasps with point contacts on complex object shapes is a particular aspect of this larger problem for which no efficient complete solutions have been found yet. In this work, we propose an efficient method that generates and locally optimizes stable force-closed grasps. Our method can be seen as complementary to state-of-the art heuristic approaches such as [3], [6]–[12].

Most robot object interactions, static or dynamic in nature, require that an object is firmly held. The force and contact position based grasp description [13], [14] is currently the most mature approach available to formalize this concept. In this work, we consider point contact dependent aspects of the grasp synthesis problem and mostly neglect other issues such as physical constraints and kinematic limitations. We furthermore utilize a simplified concept of reachability which our system integrates with a notion of grasp quality. We frame the search for force-closed grasps on 3D mesh objects as a combinatorial optimization problem and employ an iterative abstraction and refinement approach. Following the generic multilevel refinement paradigm, our approach creates a hierarchy of approximations to the grasp search problem on the original object. For this purpose, we recursively simplify the original mesh representation, resulting in a set of object models of varying resolution. We then consider the midpoint of each triangle face as a possible contact position. At each level of the hierarchy, an initial grasp is iteratively refined by local search and finally extended to the level below by finding a similar face in the higher resolution mesh. This process is exemplified in Fig. 1. The main contributions of this work can be summarized as follows:

- We introduce an approach operating directly on an input mesh, without the need for further high-level features [11], [15].
- Our approach can process complicated object geometry as well as complex and high resolution meshes consisting of thousands of faces to produce high quality grasps.
- We do not require that the input data has the same resolution at all parts of the object—a potential we leave for future exploration.
- The multilevel approach iteratively simplifies the objective function resulting in fast convergence. Furthermore, the local search at each level effectively reduces the total number of evaluated candidate solutions to a practical amount.

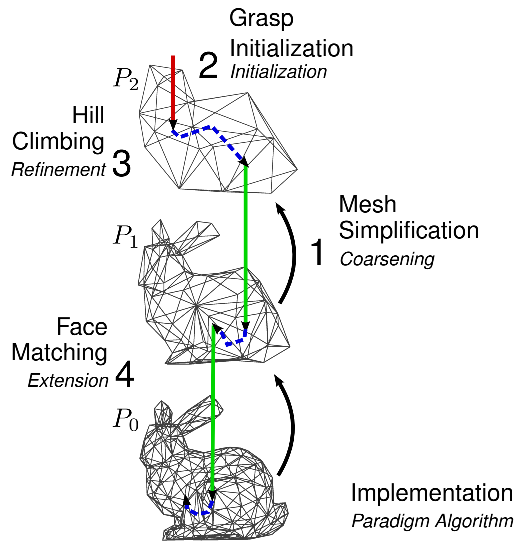

Fig. 1. System outline: Our approach applies the multilevel refinement paradigm to point contact grasps generation. We recursively simplify a mesh representation (1) and find locally optimal grasps at each level (3), starting at the topmost level (2). Solutions are transferred to the level below by finding similar faces (4). The process is sketched for one contact point only.

II. Related Work

The synthesis of stable grasps on shapes of high complexity and with exact contact positions is a challenging problem. Research has addressed grasp synthesis in various ways

The authors are with the Computer Vision and Active Perception Lab, Centre for Autonomous Systems, School of Computer Science and Communication, KTH Royal Institute of Technology, Stockholm, Sweden, {kaiyuh, jastork, fpokorny, dani}@kth.se. This work was supported by FLEXBOT (FP7-ERC-279933), the Swedish Research Council and the Swedish Foundation for Strategic Research.

which can be summarized into two main groups: Firstly, by developing compact and expressive object representations to deal with the underlying complexity of the problem space [3], [11], [15], [16], and secondly, by devising methods for analyzing and evaluating grasp quality [13], [17], [18] - in each case possibly taking into account constraints related to the task or embodiment. In this section, we discuss the two areas in relation to our contributions.

A. Object Representations for Grasp Synthesis

Object representations for grasp synthesis are often formed by extracting either global or local features from an object description. Using local features such as surface geometry properties, it is possible to evaluate grasp stability [17] and sensitivity [19], or to compute independent contact regions [20]. By exploring global features of the target object in the form of skeletal or topological descriptors, one can furthermore synthesize grasps by sampling grasp parameters such as pre-shapes, positions and approach directions of a robot hand. The reported results using these methods show relatively good success rates for obtaining feasible grasps [11], [21]. However, many of these approaches need to execute a grasp policy [6], [7] in simulation to subsequently conduct a force-analysis of the resulting grasp configuration. Therefore they can be regarded as a form of heuristic to generate point contact grasps.

Many proposed object representations extract skeletal features from 3D data [11], [15]. The Reeb Graph [11] and Medial Axis [21], were successfully used for generating initial positions, approach directions and pre-shapes for whole hand grasps [10], [11]. Similarly, the work of [22] detects holes on objects as a basis for executing caging grasps. Approximating objects' shape using basic primitives has been recognized as another object representation [3], [12], where approach vectors and pre-shapes are generated by sampling on shape primitives. By parameterizing objects using superquadrics [6], [8], [9] and superellipsoids [23], grasp parameters are sampled on the approximated objects.

Our approach relates to the above methods relying on local features since it is based on grasp quality evaluation and requires information about the surface and its normals. It is also related to methods using global features since recursively coarsening the mesh model results in an increasingly rougher object description, maintaining only global characteristics. In contrast to global feature methods, our approach optimizes grasp contacts directly without the detour of describing the grasp by grasp parameters. As a consequence, we do not require the execution of a grasp policy [6], [7] in simulation to get access to a candidate solution. Additionally, as we provide exact grasping contact positions which are required in some applications such as finger gaiting [24], [25], our method can act as a plug-in in such applications to complete the grasp synthesis step in a pipeline.

B. Grasp Contact Synthesis

Sampling-based grasp synthesis methods, as discussed above, are nonconstructive in the sense that they assess grasp stability based on contacts resulting from the execution of a grasp policy, but they do not explicitly relate the synthesis to the quality measure and solve the grasp synthesis problem indirectly. In [17], it was shown that randomized grasp generation is fast and suitable for some simple test objects. However, this is not directly applicable for objects of complex shapes as we shall show later in this paper. As also have shown in [19], sampling based methods furthermore suffer in the case of low friction coefficients, where stable grasps become increasingly sparse.

In the work of [26], a local gradient based grasp quality optimization was performed and grasps were also transferred between objects by a continuous shape/grasp deformation and optimization process. However, this approach required a particular smooth surface parametrization. In our work, the grasp configuration space is instead discretized according to a hierarchically coarsened mesh representation of the object surface. This enables an optimization of contact based grasps by a discrete local search and following a multilevel combinatorial optimization approach.

III. Preliminaries

We start by presenting a combinatorial optimization problem formalism and continue with a discussion of grasp stability and reachability metrics.

A. Combinatorial Optimization and Multilevel Refinement

Presented with a combinatorial optimization problem, where an optimum over a discrete set of candidate solutions needs to be determined, it is often infeasible to carry out an exhaustive search to determine the global optimum. In this case, the problem can be relaxed to finding a *good* solution in *reasonable* time by applying a search heuristic. The multilevel refinement paradigm is a metaheuristic that can be applied to difficult combinatorial optimization problems and describes a recursive abstraction and refinement pattern. It has been used to address a variety of problems in graph theory and numerical algebra as described in the overview work [27]–[29].

As a metaheuristic, the multilevel refinement paradigm describes a general search strategy: a problem instance is recursively approximated to form a hierarchy of increasingly approximated problem instances. Starting at the top level, a solution is found for the current instance and then extended to the level below. Provided with an initial candidate solution on the top level, the procedure finally results in a solution to the original problem defined in the lowest level.

Formally, we denote the initial problem instance by P_0, its set of candidate solutions by $\mathcal{X}_0$, and the objective function for all levels by θ. We write $P_0, P_1, \dots$ for the problem hierarchy, where each instance P_l is created recursively by a coarsening of its parent P_{l-1}. At level l, the initial solution C_l^0 is refined to C_l, where on the top level C_l^0 is provided by an initialization algorithm and for all other levels by extending $C_{l+1} \in \mathcal{X}_{l+1}$ from the child problem. A detailed description of the approach is reported in [29] and our implementation is described in Alg. 1.

Algorithm 1 Description of the multilevel refinement paradigm metaheuristic for combinatorial optimization.

Input: Problem instance P_0
Output: Solution C_0

```
l ← 0
while coarsening do
    P_{l+1} = coarsen(P_l)
    l ← l + 1
end while
C_l^0 = initialize(P_l)
C_l = refine(C_l^0, P_l)
while l > 0 do
    l ← l − 1
    C_l^0 = extend(C_{l+1}, P_l)
    C_l = refine(C_l^0, P_l)
end while
```

While the use of sophisticated techniques for refinement have been demonstrated [30]–[33], the refinement algorithm is often selected from a family of local search heuristics. There, only a small region around the current solution is explored for improvement. This notion is formalized by the neighborhood set $N(x) \subseteq \mathcal{X}_l$ of a candidate solution $x \in \mathcal{X}_l$ which defines the locally searched space.

For application, the paradigm requires problem specific definitions of algorithms for *refinement*, problem *coarsening*, *initialization*, and solution *extension*. We describe the objective function, candidate solution space and the algorithms for our grasp synthesis problem in Sec. IV.

B. Grasp Stability Metric

Many approaches to robotic grasping of rigid objects are based on force analysis and in particular on the concept of force-closure [13], [14]. To identify stable grasp configurations, the forces exerted by the robot end effector and friction between the robot hand and the object surface are considered. We choose to evaluate the force-closure property of a grasp with the L^1 grasp quality measure Q_μ reported in [17] and which is based on the Coulomb friction model.

To this end, a grasp g is formalized by m point contacts $p_1, p_2, \ldots, p_m \in \mathbb{R}^3$ and their inward-pointing unit surface normals $n_1, n_2, \ldots, n_m \in \mathbb{R}^3$. The grasp quality Q_μ is then a function of all contact positions and normals, the center of mass of the object $z \in \mathbb{R}^3$, and the friction coefficient $\mu \in \mathbb{R}^+$. A grasp is force-closed if Q_μ is larger than zero. In the following, the relation between a grasp g and z, n_i, p_i will be implicit in cases where this does not cause confusion.

C. Reachability Measure

It is possible that a given set of contacts on an object resemble a force-closed grasp, but that the robot hand is incapable of realizing this grasp. The reasons for this are several: kinematic infeasibility of the hand pose, object or self-collisions, or the required number of contact points cannot be achieved due to lack of independent joints. These issues are in part addressed in recent work [34].

In the simplest case, the individual point contacts are just too distant from each other even when kinematics and collisions are neglected. We adopt this simplified distance-based concept of reachability to approximate the robot hand's workspace. Formally, we define reachability $R(g)$ of a point contact grasp g, given by its contact positions as described in Sec. III-B, as

$$R(g) = \frac{1}{m} \sum_{i=1}^{m} \|p_i - \psi\| \tag{1}$$

where ψ is the centroid of all contact points $p_1, p_2, \ldots, p_m$ of the grasp g and $\|\cdot\|$ is the Euclidean distance. We consider a grasp as reachable if $R(g) \leq r$ for a predefined reachability upper bound $r \in \mathbb{R}^+$. Intuitively, the pairwise distances are limited by $2r$.

IV. Methodology

In this section, we describe our approach to search for grasp contacts. In Sec. IV-A, we formalize the grasp search problem on a surface mesh using the terminology of combinatorial optimization introduced in Sec. III. Subsequently we explain our realization of the multilevel paradigm algorithm.

A. Problem Formalization

To search for a stable force-closed point contact grasp on a rigid object, we represent the object's surface by a mesh of oriented triangle faces $F = \{f_1, f_2, \ldots, f_k\}$. A triangle mesh is the simplest piece-wise linear approximation of the original surface, assuming that the vertices originated from the original surface. Instead of considering the space of all possible contact positions on all the faces, we only consider face midpoints $\Delta = \{\delta_1, \delta_2, \ldots, \delta_k\}$ as possible contacts. The normal for contact point δ_i is the face normal n_i. To simplify the notation, we will refer to faces f_i and midpoints δ_i indiscriminately and let each of the symbols f_i, δ_i, or n_i refer to any subset of $\{f_i, \delta_i, n_i\}$ where the context is clear.

The previously introduced discretization induces a combinatorial search space $\mathcal{F} = \prod_{k=1}^{m} F$ where each of the m contacts has to be assigned one face or midpoint. To complete the combinatorial optimization problem, we define an objective function $\theta \colon \mathcal{F} \to \mathbb{R}$ to evaluate candidate solution by employing the grasp quality measure Q_μ described in Sec.III-B and the reachability measure R defined in Eq. (1):

$$\theta(g) = \begin{cases} Q_\mu(g) - \left(\mathrm{e}^{\alpha\left(R(g)-r\right)} - 1\right), & R(g) > r \\ Q_\mu(g), & \text{else} \end{cases} \tag{2}$$

where $\alpha \in \mathbb{R}^+$ is a penalty factor. If the contacts are reachable, θ only describes grasp stability. Unreachable grasps with $(R(g) - r) > 0$ are increasingly dominated by the negative right hand term. The further apart the contacts are positioned in space, the lower is the value of $\theta(g)$.

Some properties of the search space $\mathcal{F}$ and the objective function θ can be inspected in Fig. 2 for three contact grasps. Faces are colored by the objective function while iterating one contact over all faces and keeping the other contacts fixed. Even if only a single contact would be optimized at a time, the search space exhibits multiple local maxima at different places (red) rendering global joint optimization a

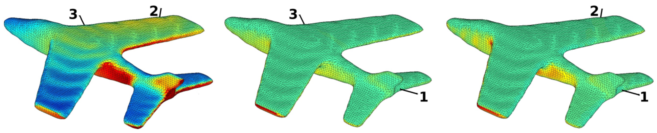

Fig. 2. Sketch of the objective function values with large reachability bound if two contacts are kept fixed. Red color marks high quality. For each contact there are multiple local maxima rendering the global joint optimization problem difficult. However, the objective function changes gradually, allowing for improvement by local search.

difficult problem. However, the objective function changes gradually from face to face which allows for exploitation with local search and multilevel refinement.

Applying the terms introduced in Sec. III-A, the search space of the initial problem instance P_0 is identified with the Cartesian product of the original mesh faces $\mathcal{X}_0 = \mathcal{F}$. Writing $\mathcal{F}_l$ and F_l, we will use subscript to refer to different levels. A problem instance P_l is coarsened to P_{l+1} by reducing the number of faces in F_l using a mesh simplifying procedure described in Sec. IV-B. We stop coarsening the mesh before it degenerates and then apply a random initialization of contacts. For the refinement step, we apply a greedy hill climbing procedure explained in Sec. IV-C using the Cartesian product of the η^{th}-order neighbor faces of a contact to form the search neighborhood. A level l candidate solution $C_l = \left(f_{i_k}\right)_{k=1}^{m} \in \mathcal{F}_l$ is extended to level $l-1$ by matching $f_{l_i} \in C_l$ to a similar face in $\mathcal{F}_{l-1}$, as explained in Sec. IV-D.

B. Mesh Simplification

For the *coarsening* procedure, we automatically produce simpler and approximated versions of mesh models F_l using the surface simplification algorithm of [35]. The algorithm was suggested for multi-resolution modeling and iteratively produces high quality approximations of polygonal models at a decreasing level of detail. Each simplification step joins one pair of vertices which is selected according to a per-vertex surface error metric. In the contraction, all incident edges are connected to the remaining vertex, degenerated edges and faces are removed, and the vertex is placed to minimize a quadric surface error metric. This procedure ensures high fidelity to the original model in every simplification step.

Since the algorithm removes only one vertex per iteration, a sequence of models with only gradually decreasing level of detail is produced in place. We apply a percentage reduction of the number of faces as termination condition for each recursive coarsening step. Thereby, we decide the number of levels $u \in \mathbb{N}$ and final number of faces $o \in \mathbb{N}$ beforehand:

$$|F_0|\,\gamma^u = o \qquad \text{and} \qquad |F_l|\,\gamma = |F_{l+1}| \tag{3}$$

In this way, we obtain a sequence of models having the same ratio of $1-\gamma$ less vertices with respect to the previous one. An example of such a sequence of mesh models is depicted in Fig. 1.

C. Hill Climbing

For the *refinement* procedure, we turn to local search and employ hill climbing until convergence. Given an initial candidate solution $C_l^0 = \left(f_{i_k}\right)_{k=1}^{m} \in \mathcal{F}_l$, we define the neighborhood of C_l^0 as the Cartesian product of the η^{th}-order neighboring faces for each $f_i \in C_l^0$. Formally, we write

$$N(C_l^0) = \prod_{k=1}^{m} N^\eta(f_{k_1}), \tag{4}$$

where $N^\eta(f_i)$ denotes the set of triangles in the mesh F_l which can be reached from f_i by η or less steps, and including f_i itself. An example of neighbor faces at different step distances are shown in the left of figure Fig. 3.

In each hill climbing iteration, we select the best grasp from $N(C_l)$ until no improvement is achieved. This procedure is formalized in Alg. 2. We maintain a look-up table for already calculated grasps to reduce time complexity. In experiments, we have found that it is particularly beneficial to increase η from 1 to 2 when the current grasp is unstable. This technique is later referred to as an *adaptive* approach.

Algorithm 2 Description of the refinement procedure using hill climbing until convergence.

Input: Initial grasp $C_l^0 = \left(f_{i_k}\right)_{k=1}^{m} \in \mathcal{F}_l$
Output: Solution C_l
  $i \leftarrow 0$
  repeat
    $N = N(C_l^i)$
    $i \leftarrow i + 1$
    $C_l^i = \underset{c \in N}{\operatorname{argmax}} \left(\theta\left(c\right)\right)$
  until $C_l^i = C_l^{i-1}$
  $C_l \leftarrow C_l^i$

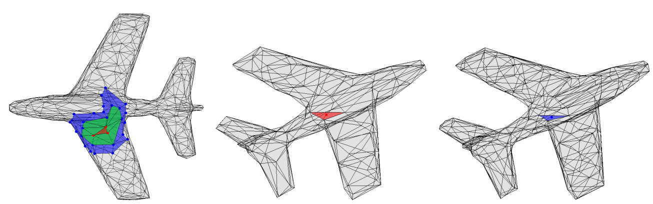

Fig. 3. **Left:** The neighboring faces in 0, 1, and 2 steps distance in red, green, and blue color respectively. The Cartesian product of each contact's neighbor faces up to η steps distance forms the neighborhood of a grasp. **Right:** Matching a face of a low resolution mesh (red) to a high resolution mesh (blue) using 6D coordinates comparison.

D. Face Matching

The *extension* procedure in our approach translates a grasp $C_l = \left(f_{i_k}\right)_{k=1}^{m}$ from the mesh F_l to the mesh F_{l-1}. Since the objective function heavily depends on the contact positions and contact normals, we match every face $f_i \in C_l$ with a face in F_{l-1} that has similar midpoint and normal. For this, we introduce a distance function between pairs of faces. First, we translate and scale each mesh F_l such that the centroid is in the origin and each vertex is at most at distance 1 from the origin. This leads to normalized midpoint positions $\bar{\delta}_i$. A face $f_i \in F_l$ is matched with a face $f \in F_{l-1}$ according to the following

$$f = \underset{f_k \in F_{l-1}}{\operatorname{argmin}} \left(\left\| (\bar{\delta}_i, n_i) - (\bar{\delta}_k, n_k) \right\| \right), \tag{5}$$

where $(\bar{\delta}_i, n_i) \in \mathbb{R}^6$. Equation (5) can be efficiently implemented using a k-d tree data structure [36]. Fig. 3 shows how a face on a coarser mesh (red) is matched to a face (blue) on a finer mesh on the right side.

V. Experiments

In this section, we present empirical evidence for the viability of our approach. For all experiments, the Coulomb friction coefficient is set to $\mu = 1$ and the penalty factor in the objective function is $\alpha = 4$. The models *bunny*, *plane*, *dinopet*, and *homer* are found in [37], [38] and have 52.000, 17.354, 8996, and 10.202 faces respectively. Each model was centered and scaled so that the maximum distance between any vertex and the origin is 1. Unless stated otherwise, the top level always has 100 faces, we use 4 coarsening levels resulting in 5 levels in total, and the number of contacts is $m = 3$. Figures 5, 7, 9, and 11 show standard box-plots indicating median, 25th and 75th percentiles and individual outliers.

A. Varying the Number of Levels

First, we investigate how the number of recursive coarsening steps influences the quality and reachability of the resulting grasp. For each model, we consider 0 to 4, and 6, 8, and 10 coarsening steps with 100 faces at the top level. We generate 100 random initial grasps on the coarsest mesh and execute our approach for each of these 8 settings. The reachability bound is always set to $r = 0.5$ and only the 1st-order neighborhood is considered.

The ratio of reachable grasps is described in Fig. 4. A maximum is generally reached for level 6 or 8, but in each setup at least 73% of the grasps were reachable. For the models *bunny*, *plane*, and *homer*, the percentage of reachable grasps was over 90% for more than 2 coarsening levels. Since the initialization was done randomly and our approach performs greedy optimization, it was not always possible to move the contacts close enough to satisfy the reachability constraint, particularly for the highly concave object *dinopet*.

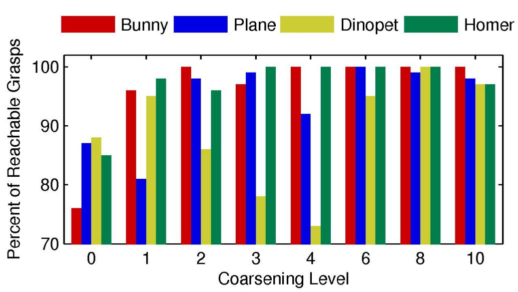

Fig. 4. The percentage of reachable grasps for different numbers of levels. Each time, 100 executions with random initialization were considered.

The distribution of grasp quality Q_μ for reachable grasps can be inspected in Fig. 5. We observe a general trend showing that grasp quality is increased and variance reduced as the number of levels is increased. However, one recursive coarsening step already improves results considerably over mere hill climbing on the original mesh at level 0.

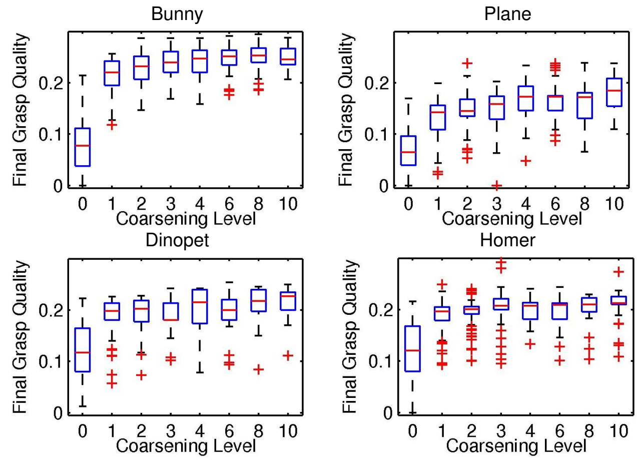

Fig. 5. Grasp quality distribution for the reachable grasps from 100 runs with random initialization. Generally, grasp quality improves with the number of levels used.

B. Empirical Complexity

Since the experiment in the previous section shows that more levels in general lead to better expected grasp quality, we have to investigate how the average and worst case complexity relate to the number of levels. For this, we consider how many hill climbing steps the approach requires at each level and how many grasps need to be compared in each of these steps. The presented data is taken from the above experiment and for the *bunny* object.

The average and maximal number of steps per level in relation to the number of coarsening steps is depicted in Fig. 6. When we initialized the algorithm with 4 or more levels, the maximal number of steps per level were similar and initialization with 0–3 levels required substantially more steps. For 2 or more initialization levels, the average number of steps per level stabilized between 2.5 and 3.1, and 3.4 steps were required in the final level. However, the number of grasp quality evaluations per hill climbing step depends on the cardinality of the set of neighbor faces. A statistic for the 1^{st}-order neighbors on the *bunny* model for different coarsening levels is shown in Fig. 7. In our experiment, the average number of 1^{st}-order neighbors did not change for different levels and so the complexity of each hill climbing step is not influenced by the number of levels.

		Max. #Steps								**Avg. #Steps**							
	Level	0	1	2	3	4	6	8	10	0	1	2	3	4	6	8	10
Initial Level	10	5	5	4	5	7	5	5	6	2.6	2.6	2.5	2.5	2.7	2.1	2.4	3.4
	8	6	6	5	6	6	4	6		3.0	2.8	2.2	3.1	2.9	2.6	3.4	
	6	4	4	5	6	5	6			2.6	2.7	2.6	2.5	2.5	3.4		
	4	6	5	7	5	6				3.0	3.0	3.0	2.6	3.4			
	3	7	10	8	6					3.0	3.4	2.8	3.4				
	2	6	7	6						3.3	3.1	3.4					
	1	16	6							4.7	3.4						
	0	38								10.6							

Fig. 6. Maximum and average number of hill climbing steps sorted by initial level and for the experiment in Sec. V-A for the *bunny* model.

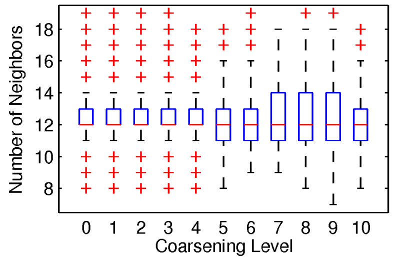

Fig. 7. Statistics of the number of 1^{st} order neighbor faces for the *bunny* model for different coarsening levels. The average number of faces is stable and so the complexity of one hill climbing step does not depend on the number of levels.

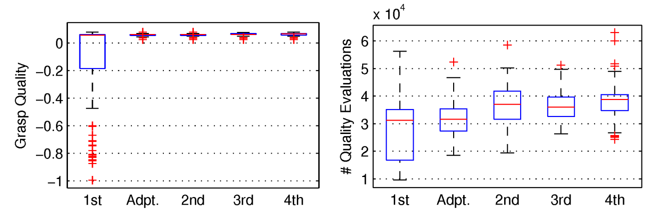

Fig. 9. Comparison of the final grasp quality and number of grasp quality evaluations when the hill climbing procedure considered only 1^{st}-order, adaptively 1^{st} or 2^{nd}-order (Adpt.), 2^{nd}-order, 3^{rd}-order, or 4^{th}-order neighborhoods.

C. Complexity of Objects

Intuitively, grasp synthesis is easier on simpler objects. We investigate this claim in Fig. 8, where grasp quality distributions for objects of varying complexity are displayed. The setup of this experiment is the same as in Sec. V-A except that the reachability bound was set to $r = 1.0$ and only 4 coarsening levels were used. In our experiment, the grasp quality variance increases with the complexity of the object, since the number of local optima increases with the complexity of the object.

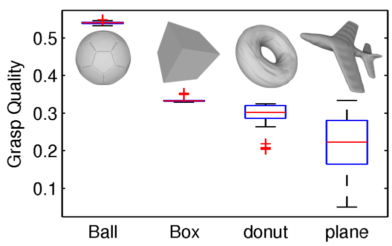

Fig. 8. Grasp quality distribution for objects of varying complexity.

D. Neighborhood Orders

In the previous experiment, we saw that the level number does not dramatically influence the average amount of 1st-order neighboring faces which governs the complexity of a hill climbing step as described in Eq. (4). According to Sec. IV-C it is possible to consider a larger neighborhood for local search or even to adapt the order of the neighborhood dynamically. We now investigate the influence of the size of the neighborhood on the final grasp quality and on the number of grasp quality evaluations. We generate 100 random initial grasps for the *plane* model and compare the results for different sizes of neighborhoods. The reachability bound is set to a small value $r = 0.133$ and we compare 1^{st}-order 2^{nd}-order, 3^{rd}-order, and 4^{th}-order neighborhoods, as well the adaptive approach mentioned in Sec. IV-C that increases η from 1 to 2 if the current grasp is not force-closed.

The results presented in Fig. 9 show that the low reachability bound leads to many unstable grasps for the 1^{st}-order approach. Variance in grasp quality reduces if more than the 1^{st}-order neighborhood is considered. This means that only considering 1^{st}-order neighborhoods makes the method more sensitive to local maxima. Using larger neighborhoods comes at a cost in form of an increased number of grasp quality evaluations. However, the number of grasp quality evaluations is not dramatically increased, since – when a larger neighborhood is considered – the algorithm tends to converge much quicker and requires less optimization steps at each level.

E. Comparison to Random Sampling

From the previous experiment, we know that our approach requires less than 40.000 grasp quality evaluations on average on the objects considered. We now compare the results of our adaptive method using random initialization to a batch sampling approach where the best of 50.000 reachable randomly sampled grasps from $\mathcal{F}_0$ is selected in each batch. We set a small reachability bound $r = 0.133$ and executed 100 runs each for both methods on the model *plane*.

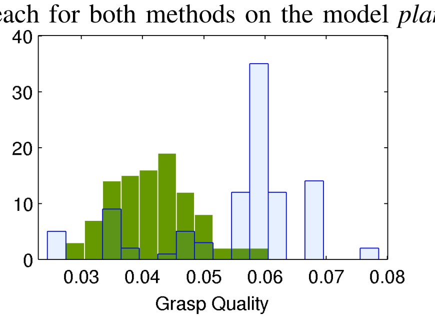

Fig. 10. Histogram of maximal grasp quality from randomly sampled reachable face midpoints on the original mesh (green) and grasp quality when our method with random initialization, adaptive neighborhood size and 4 coarsening steps (blue) is used on the *plane* model.

Grasp quality results are presented in Fig. 10 where it can be clearly seen that the best grasps of each batch were on average worse than the resulting grasps from our approach. The low overall value of grasp quality is to be attributed to the small reachability bound which reduces the quality measure. While our approach stayed below one minute total execution time, the sampling-based approach consumed 85s in our implementation.

F. Heuristic Initialization

In all previous experiments, we used a random initialization for our approach. However, our approach is designed to improve grasp quality and therefore can be complementary to any heuristic that generates initial grasp contacts. In this experiment, we compare the performance of our approach to a simplified automatic fingertip closing grasp policy and initialize both using the same heuristic.

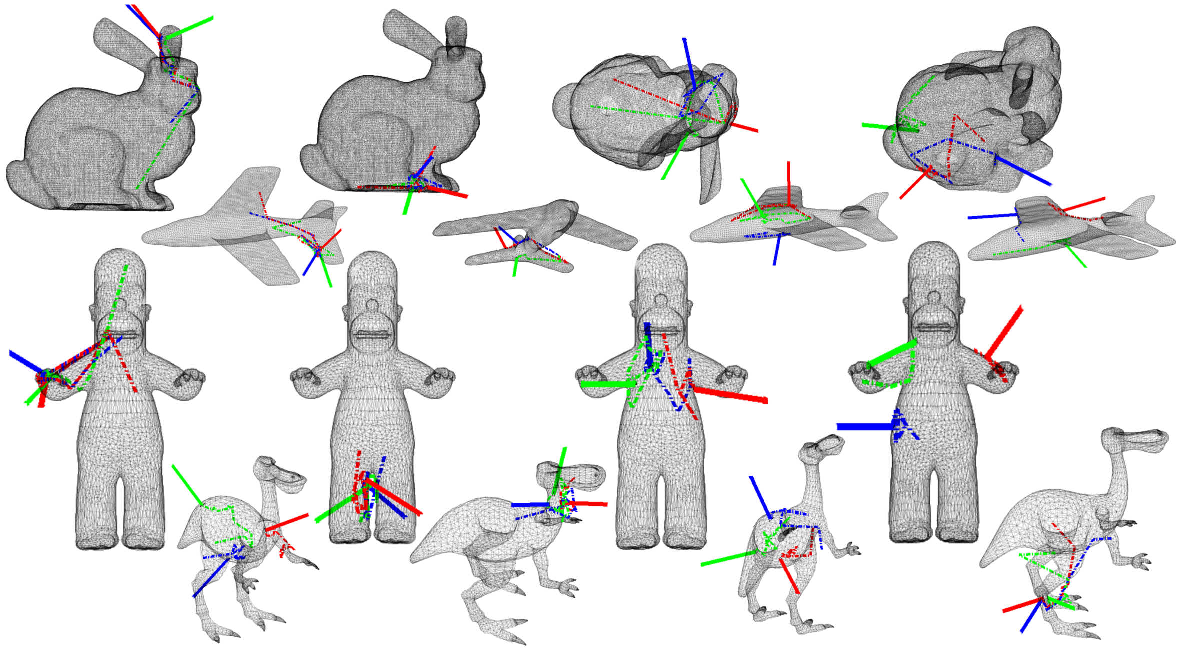

Fig. 12. A couple of examples from the evaluation runs: Provided random initialization, our approach is capable to synthesize point contact grasps that comply to the different reachability constraints. The dotted lines indicate the path each contact took in 3D space during the iterated refinement and hill climbing steps. In many cases, the contact positions had to be heavily adjusted to fulfill reachability. The reachability bounds used in the example were (from left to right): r = 0.033, 0.066, 0.2 and 0.33.

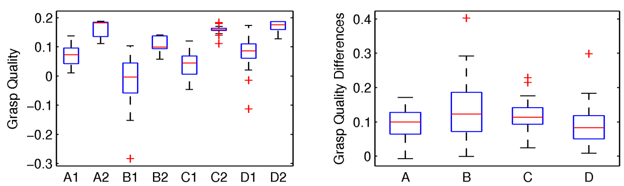

Fig. 11. Comparison between automatic closing of fingertips and our approach when both are initialized with a heuristic. The used mesh models are *bunny* (A), *plane* (B), *dinopet* (D), and *homer* (D). Fingertip closing is indicated by (1) and our approach is marked with (2).

The initialization heuristic used, considers the same set of vectors $(\bar{\delta}_i, n_i) \in \mathbb{R}^6$ as described in Sec. IV-D. We cluster the data $\{(\bar{\delta}_i, n_i) \mid f_i \in F_4\}$ into 6 sets $S_1, S_3, \ldots, S_6$ using k-means and consider each combinations of clusters $\{S_i, S_j, S_l\}$. From each combination, we retain the best quality grasp, resulting in 20 different grasps. These grasps serve directly as initialization for our approach. Our automatic closing considers the contact position and normals as grasp parameters. For each contact, a virtual fingertip is placed on an approach line outside of the object. Now using the original mesh model F_0, all fingertips are moved along the approach line until collision, where contact positions and normals are recorded. The reachability bound $r = 0.33$ was used for this experiment.

The results depicted in Fig. 11 suggest that automatic fingertip closing generally produces grasps with lower quality. On the right-hand side of Fig. 11, we show the difference in grasp quality for the same initial grasp. In the majority of the cases, the difference was positive, indicating that our approach performed better for almost every initialization.

G. Reachability Bound and Number of Contacts

Finally, we present qualitative results that support our choice of reachability measure explained in Sec. III-C. For a reachability bound of 0.033, 0.066, 0.2, and 0.33, we execute our adaptive approach for several mesh models. The results are depicted in Fig. 12 where it can be seen that despite random initialization, our approach is capable to synthesize point contact grasps that comply to different reachability constraints. The dotted lines indicate the path each contact took in 3D space during the iterated refinement and hill climbing steps. In many cases the contact positions had to be heavily adjusted to fulfill reachability. By setting reachability bound to $r = 0.33$, we also show that our approach is able to synthesize point contact grasps with varying numbers of contacts. As is to be expected, Fig. 13 shows that grasp quality is improved the more contacts are available.

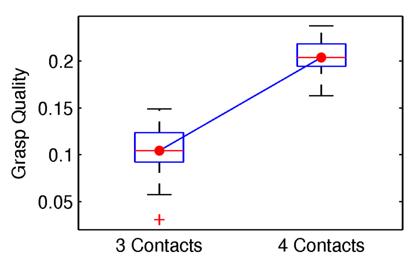

Fig. 13. Grasp quality distribution for synthesized grasps on the *plane* model with 3 and 4 contacts.

VI. Conclusions

We proposed a method for synthesizing grasps based on point contact search using search space discretization. Grasp synthesis is formulated as a combinatorial optimization problem. Based on a multilevel refinement, we have considered a hierarchy of recursively coarsened mesh models and locally improved grasps at each level. For this, we have optimized grasp quality subject to a reachability measure at each level and matched grasps to the next lower level, to finally achieve a grasp on the original object model.

Our approach is applicable to complex input meshes, having thousands of faces and highly concave shape. Thus, it provides the first tractable method for search of grasp contacts on such input data. Our empirical evaluation shows that the method produces feasible, high quality grasps from random and heuristic initializations. It outperformed random sampling relying on a large number of generated grasps and automatic hand closing technique. Analyzing the influence of object model properties and method parameters such as the number of coarsening levels and the size of the local search space, we have shown that our approach is viable. In the future, we plan to additionally consider hand kinematics in the optimization function and we intend to study the influence of input data uncertainty on our approach.

VII. Acknowledgement

We thank Manuela Ortlieb, Research Group Foundations of AI, University of Freiburg, Germany, for her useful suggestions.

References

[1] Z. Liu, L. B. Gueta, and J. Ota, "A strategy for fast grasping of unknown objects using partial shape information from range sensors," *Advanced Robotics*, vol. 27, no. 8, pp. 581–595, 2013.

[2] M. Popovic, D. Kraft, L. Bodenhagen, E. Baeski, N. Pugeault, D. Kragic, T. Asfour, and N. Krüger, "A strategy for grasping unknown objects based on co-planarity and colour information," *Robotics and Autonomous Systems*, vol. 58, no. 5, pp. 551 – 565, 2010.

[3] K. Huebner, "BADGrA toolbox for box-based approximation, decomposition and GRasping," *Robotics and Autonomous Systems*, vol. 60, no. 3, pp. 367 – 376, 2012.

[4] Y. Bekiroglu, D. Song, L. Wang, and D. Kragic, "A probabilistic framework for task-oriented grasp stability assessment," in *IEEE ICRA*, Karlsruhe, Germany, May 6–10 2013.

[5] M. Gabiccini and A. Bicchi, "On the role of hand synergies in the optimal choice of grasping forces," in *Robotics: Science and Systems*, Zaragoza, Spain, June 2010.

[6] D. Berenson, R. Diankov, K. Nishiwaki, S. Kagami, and J. Kuffner, "Grasp planning in complex scenes," in *IEEE-RAS Humanoids*, 2007, pp. 42–48.

[7] A. Miller and P. Allen, "Graspit! a versatile simulator for robotic grasping," *Robotics Automation Magazine, IEEE*, vol. 11, no. 4, pp. 110–122, 2004.

[8] G. Biegelbauer and M. Vincze, "Efficient 3d object detection by fitting superquadrics to range image data for robot's object manipulation," in *IEEE ICRA*, 2007, pp. 1086–1091.

[9] T. T. Cocias, S. M. Grigorescu, and F. Moldoveanu, "Multiple-superquadrics based object surface estimation for grasping in service robotics," in *Optimization of Electrical and Electronic Equipment (OPTIM)*, 2012, pp. 1471–1477.

[10] N. Vahrenkamp, M. Przybylski, T. Asfour, and R. Dillmann, "Bimanual grasp planning," in *IEEE-RAS Humanoids*, 2011, pp. 493–499.

[11] J. Aleotti and S. Caselli, "A 3d shape segmentation approach for robot grasping by parts," *Robot. Auton. Syst.*, vol. 60(3), pp. 358–366, 2012.

[12] A. T. Miller, S. Knoop, H. I. Christensen, and P. K. Allen, "Automatic grasp planning using shape primitives," in *IEEE ICRA*, vol. 2, 2003, pp. 1824–1829.

[13] C. Ferrari and J. Canny, "Planning optimal grasps," in *Robotics and Automation, Proceedings.* IEEE, 1992, pp. 2290–2295.

[14] A. Bicchi and V. Kumar, "Robotic grasping and contact: A review," in *IEEE ICRA*, vol. 1, 2000, pp. 348–353.

[15] M. Przybylski, T. Asfour, and R. Dillmann, "Unions of balls for shape approximation in robot grasping," in *IEEE/RSJ IROS*. IEEE, 2010, pp. 1592–1599.

[16] A. Sahbani, S. El-Khoury, and P. Bidaud, "An Overview of 3D Object Grasp Synthesis Algorithms," *Robotics and Autonomous Systems*, vol. 60, no. 3, pp. 326–336, 2011.

[17] C. Borst, M. Fischer, and G. Hirzinger, "Grasping the dice by dicing the grasp," in *IEEE/RSJ IROS*, vol. 4, 2003, pp. 3692–3697 vol.3.

[18] F. T. Pokorny and D. Kragic, "Classical grasp quality evaluation: New theory and algorithms," in *IEEE/RSJ IROS*, Tokyo, Japan, 2013.

[19] K. Hang, F. T. Pokorny, and D. Kragic, "Friction coefficients and grasp synthesis," in *IEEE/RSJ IROS*, Tokyo, Japan, 2013.

[20] M. Roa and R. Suarez, "Computation of independent contact regions for grasping 3-d objects," *Robotics, IEEE Transactions on*, vol. 25, no. 4, pp. 839–850, 2009.

[21] T. K. Dey, "Approximate medial axis as a voronoi subcomplex," in *In Proc. 7th ACM Sympos. Solid Modeling Applications*. ACM Press, 2002, pp. 356–366.

[22] F. T. Pokorny, J. A. Stork, and D. Kragic, "Grasping objects with holes: A topological approach," in *IEEE ICRA*, 2013.

[23] R. Pelossof, A. Miller, P. Allen, and T. Jebara, "An svm learning approach to robotic grasping," in *IEEE ICRA*, vol. 4, 2004, pp. 3512–3518 Vol.4.

[24] J.-P. Saut, A. Sahbani, S. El-Khoury, and V. Perdereau, "Dexterous manipulation planning using probabilistic roadmaps in continuous grasp subspaces," in *IEEE/RSJ IROS*, 2007, pp. 2907–2912.

[25] L. Han and J. Trinkle, "Dextrous manipulation by rolling and finger gaiting," in *IEEE ICRA*, vol. 1, 1998, pp. 730–735 vol.1.

[26] F. T. Pokorny, K. Hang, and D. Kragic, "Grasp moduli spaces," in *Proceedings of Robotics: Science and Systems*, June 2013.

[27] A. Brandt, "Multilevel computations: Review and recent developments," *Multigrid Methods: Theory, Applications and Supercomputing*, pp. 35–62, 1988.

[28] S.-H. Teng, "Coarsening, sampling, and smoothing: Elements of the multilevel method," in *Algorithms for Parallel Processing*. Springer, 1999, pp. 247–276.

[29] C. Walshaw, "Multilevel refinement for combinatorial optimisation problems," *Annals of Operations Research*, vol. 131, no. 1-4, pp. 325–372, 2004.

[30] D. Vanderstraeten, C. Farhat, P. Chen, R. Keunings, and O. Ozone, "A retrofit based methodology for the fast generation and optimization of large-scale mesh partitions: beyond the minimum interface size criterion," *Computer methods in applied mechanics and engineering*, vol. 133, no. 1, pp. 25–45, 1996.

[31] A. Kaveh and H. Bondarabady, "A hybrid graph-genetic method for domain decomposition," *Finite elements in analysis and design*, vol. 39, no. 13, pp. 1237–1247, 2003.

[32] M. Toulouse, K. Thulasiraman, and F. Glover, "Multi-level cooperative search: a new paradigm for combinatorial optimization and an application to graph partitioning," in *Euro-Par99 Parallel Processing*. Springer, 1999, pp. 533–542.

[33] A. Langham and P. Grant, "A multilevel k-way partitioning algorithm for finite element meshes using competing ant colonies," in *the Genetic and Evolutionary Computation Conf*, vol. 2, 1999, pp. 1602–1608.

[34] Z. Xue and R. Dillmann, "Efficient grasp planning with reachability analysis," *International Journal of Humanoid Robotics*, vol. 8, no. 04, pp. 761–775, 2011.

[35] M. Garland and P. S. Heckbert, "Surface simplification using quadric error metrics," in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 1997, pp. 209–216.

[36] J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Communications of the ACM*, vol. 18, no. 9, pp. 509–517, 1975.

[37] H. Benhabiles, J.-P. Vandeborre, G. Lavoué, and M. Daoudi, "A framework for the objective evaluation of segmentation algorithms using a ground-truth of human segmented 3D-models," in *IEEE SMI*, Beijing, China, June 26-28 2009.

[38] X. Chen, A. Golovinskiy, and T. Funkhouser, "A benchmark for 3D mesh segmentation," *ACM Transactions on Graphics (Proc. SIGGRAPH)*, vol. 28, no. 3, Aug. 2009.