

Advancing Robust Multi-Object Manipulation with Energy Margin

Yifei Dong[†], Xianyi Cheng[‡], Werner Friedl[§], Aurel Schröter[§], Ashok M. Sundaram[§],
Máximo A. Roa[§] and Florian T. Pokorny[†]

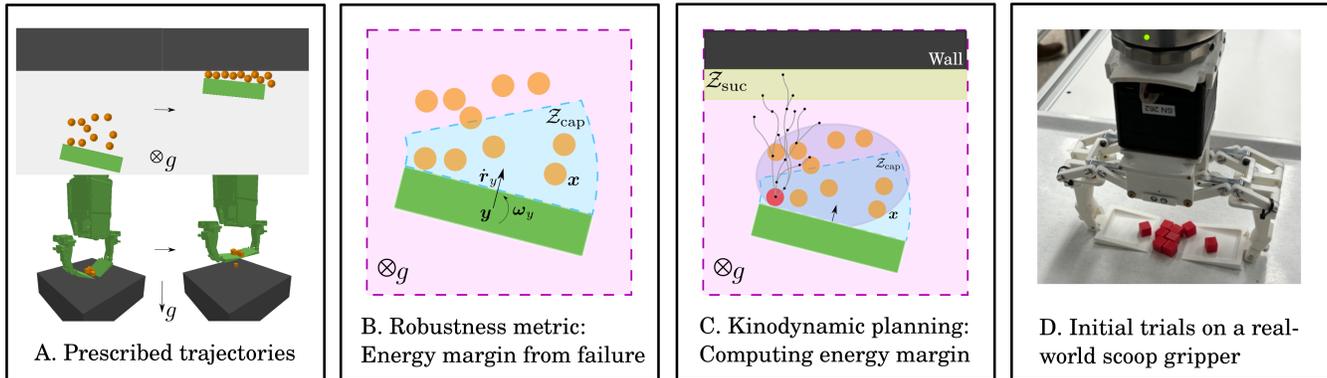


Fig. 1: Overview of the Robustness Analysis Framework. A: Simulation setup showing trajectories for pushing and scooping tasks with multiple objects (orange) and robot end-effectors (green). B: Frame-by-frame robustness analysis using energy margins to evaluate the likelihood of robust and successful manipulation. C: Calculation of energy margins involves applying serial random wrench perturbations to each object, accounting for the predictive model’s uncertainty. Potential paths for one exemplar object (red) is visualized. D: Real-world application with an initial experiment on multi-object scooping, highlighting future directions for this research.

Abstract—Quantifying robustness is crucial for developing effective manipulation policies. However, assessing the robustness of manipulating multiple objects simultaneously is challenging due to uncertainties in predictive models, complex interactions among objects, and the high-dimensional nature of the joint state space. This workshop contribution introduces an approach for analyzing manipulation robustness through energy margins and caging-based analysis. Our approach evaluates multi-object manipulation robustness by measuring the energy margin to failure. Enabled by a kinodynamic planning framework that incorporates inter-object interactions and contact dynamics, this method provides a systematic way to predict manipulation robustness and success under uncertainty. We demonstrate the utility of this approach with two simulated tasks, scooping and planar pushing, highlighting its capability to handle and represent high-dimensional state spaces effectively.

I. INTRODUCTION

Human manipulation is notable for its dexterity, simplicity, and robustness, often allowing the simultaneous manipulation of multiple objects. However, existing methods for evaluating the robustness of such multi-object manipulation are limited. The main challenges include: (1) complex inter-object interactions and nonholonomic contact constraints [1], (2) inaccuracies in predictive models due to the uncertainties in contact modeling [2], and (3) the high dimensionality of

the combined state space of objects and the robot, often leading to computational difficulties.

Current research on robustness is generally divided into three categories: grasp quality metrics [3], [4], control stability analysis (including contraction analysis [5], Lyapunov functions [6] and barrier functions [7]), and caging. Most quality metrics in grasping [8] and non-prehensile manipulation [9] [10] usually define robustness through object positional or wrench space margins based on first-order analysis. Control stability methods offer a local perspective that struggles to account for the robustness of global geometric properties. These methods above often fail to address the full complexity of multi-object manipulation, particularly when it involves complicated interactions and dynamic changes.

Caging approaches, which confine an object within a bounded space to prevent escape, provide a more global geometric analysis of robustness. Caging was introduced by Kuperberg [11] and further developed by Rimon et al. [12], [13]. Caging an object only partially has been enhanced by considering external forces such as gravity [14] or elastic forces [15]. Partial caging can thereby robustly secure an object under an external perturbation level. Caging in motion [16], which this work is established upon, expands on prior quasi-static cage [17] and soft fixture [15] analyses by adapting classical caging concepts to a kinodynamic context.

Inspired by the line of energy-based caging research [14], [18], [15] and extending [16], this workshop contribution introduces a new lens by evaluating the robustness of manipulation strategies through energy margins from failure. A sampling-based kinodynamic planning framework

[†]: The authors are with the division of Robotics, Perception and Learning, KTH Royal Institute of Technology, 10044 Stockholm, Sweden. [‡]: The author is with Carnegie Mellon University, Pittsburgh, PA, 15213, USA. [§]: The authors are the Institute of Robotics and Mechatronics, German Aerospace Center (DLR), 82234 Wessling, Germany. This work is partially funded by the European Commission under the Horizon Europe Framework Programme project SoftEnable, grant number 101070600, <https://softenable.eu/>. Contact: yifeid@kth.se

is employed for energy margin computation, designed to be inherently efficient in dynamic, contact-rich environments with multiple objects. Utilizing forward simulation enables natural consideration of inter-object contact interactions and complex geometries. We validate our approach through simulated tasks of scooping and planar pushing, demonstrating its effectiveness and scalability for handling multiple objects. Additionally, we introduce a novel scooping gripper in our experiments, showcasing its utility in simultaneous multi-object grasping tasks.

II. ENERGY MARGIN: DEFINITION AND COMPUTATION

This section introduces the concept of energy margin as a metric for evaluating the robustness of multi-object manipulation strategies. We define key terms and outline the computational methods used to quantify the energy margin within a dynamic manipulation context.

A. Nomenclature: In our framework, we define the configuration space (C-space) of a 3D rigid object as $\mathcal{X} \subset \text{SE}(3)$. The C-space for the robot's end-effector and surrounding static obstacles is denoted as $\mathcal{Y} \subset \text{SE}(3) \times \mathbb{R}^{n_r}$, where $\text{SE}(3)$ represents the base pose and n_r refers to the number of robot joints. An element $\mathbf{x} \in \mathcal{X}$ represents the position and orientation of the object, and similarly, $\mathbf{y} \in \mathcal{Y}$ represents the state of the end-effector. For multi-object manipulation involving n_o objects, we introduce a joint state space $\mathcal{Z} = \mathcal{X}^{n_o} \times \mathcal{X}^{n_o} \times \mathcal{Y} \times \mathcal{Y}$. The energy function $E(\mathbf{z})$ encapsulates the total mechanical energy of the system at \mathbf{z} .

B. Defining Energy Margin: For a system with an initial state \mathbf{z}_{init} , we first define a *capture set* $\mathcal{Z}_{\text{cap}}(\mathbf{z}_{\text{init}}) \subset \mathcal{Z}$, which considers the objects as dynamically controllable by the end-effector for achieving specific tasks. Similarly, we define a *task success set* $\mathcal{Z}_{\text{suc}} \subset \mathcal{Z}$, indicating the set of states symbolizing the objects accomplish the manipulation objective.

We consider an energy cost field and its correlating probabilistic distribution of the system. Specifically, we sample n_o random wrenches from a given probability distribution and apply them on the n_o objects at \mathbf{z}_{init} respectively. The system thus transitions to \mathbf{z}^0 with energy cost c^0 . By repeating the process M times from \mathbf{z}_{init} , the system terminates in a list of states $\{\mathbf{z}^0, \dots, \mathbf{z}^M\}$ with corresponding costs $\{c^0, \dots, c^M\}$. Data pairs in $\{(\mathbf{z}^0, c^0), \dots, (\mathbf{z}^M, c^M)\}$ constitute an energy cost field that demonstrates the state space reachability in terms of energy cost from \mathbf{z}_{init} . We can thus build a probability mass function $L : \mathcal{Z} \rightarrow \mathbb{R}^+$ from the energy cost field, approximating the likelihood of reaching a state $\mathbf{z}^m \in \mathcal{Z}$ with a softmax function,

$$L(\mathbf{z}^m) = \frac{e^{-\lambda(c^m - c_{\min})}}{\sum_{m=0}^M e^{-\lambda(c^m - c_{\min})}}, \quad (1)$$

where $c_{\min} = \min_{0 \leq m \leq M} c^m$ and λ is a hyper-parameter. The approximated likelihood function indicates some states with lower costs are probabilistically more reachable.

The *capture score* is defined as the sum of likelihood

Algorithm 1: Compute Performance Scores

```

1 Initialize,  $P_\infty, \mathcal{Z}_{\text{cap}}(\mathbf{z}_{\text{init}}), \mathcal{Z}_{\text{suc}}, M$ 
2 ▷ Obtain energy cost field by growing a tree
3  $\{(\mathbf{z}^0, c^0), \dots, (\mathbf{z}^M, c^M)\} \leftarrow \text{EST}(P_\infty, M)$ 
4 ▷ Approximate probabilistic distribution
5  $c_{\min} = \min_{0 \leq m \leq M} c^m$ 
6  $S \leftarrow \sum_{m=0}^M e^{-\lambda(c^m - c_{\min})}$ 
7 for  $m = 1, 2, \dots, M$  do
8    $L(\mathbf{z}^m) \leftarrow \frac{1}{S} e^{-\lambda(c^m - c_{\min})}$  ▷ EQ. (1).
9 ▷ Compute scores
10  $\Omega_{\text{cap}}(\mathbf{z}_{\text{init}}) \leftarrow \sum_{m=0}^M \delta(\mathbf{z}^m \in \mathcal{Z}_{\text{cap}}(\mathbf{z}_{\text{init}})) \cdot L(\mathbf{z}^m)$  ▷ EQ. (2).
11  $\Omega_{\text{suc}}(\mathbf{z}_{\text{init}}) \leftarrow \sum_{m=0}^M \delta(\mathbf{z}^m \in \mathcal{Z}_{\text{suc}}) \cdot L(\mathbf{z}^m)$  ▷ EQ. (3).
12 Return  $\Omega_{\text{cap}}(\mathbf{z}_{\text{init}}), \Omega_{\text{suc}}(\mathbf{z}_{\text{init}})$ 

```

values of samples in a capture set $\mathcal{Z}_{\text{cap}}(\mathbf{z}_{\text{init}})$,

$$\Omega_{\text{cap}}(\mathbf{z}_{\text{init}}) = \sum_{m=0}^M \delta(\mathbf{z}^m \in \mathcal{Z}_{\text{cap}}(\mathbf{z}_{\text{init}})) \cdot L(\mathbf{z}^m), \quad (2)$$

where $\delta(\cdot)$ is an indicator function that equals 1 if the condition inside the brackets is satisfied and 0 otherwise. Similarly, a *success score* is thus given by

$$\Omega_{\text{suc}}(\mathbf{z}_{\text{init}}) = \sum_{m=0}^M \delta(\mathbf{z}^m \in \mathcal{Z}_{\text{suc}}) \cdot L(\mathbf{z}^m), \quad (3)$$

which is essentially a predictor for fulfilling a task-specific objective from the state \mathbf{z}_{init} . In practice, we employ kinodynamic motion planners rather than repetitive Monte Carlo rollouts from \mathbf{z}_{init} . Sampling-based kinodynamic motion planners are in general more time- and memory-efficient by using strategies such as biased sampling and caching explored nodes.

C. Computing Energy Margin: To approximate the performance scores, we construct an energy cost field through kinodynamic tree expansion. Achieving an evenly distributed sampling of state space around the initial state \mathbf{z}_{init} is critical for this purpose. Expansive Space Tree (EST) is thereby selected for its efficacy in promoting uniform distribution across the state space through inverse density weighting, unlike Random-exploring Random Tree (RRT), which may lead to uneven exploration.

We consider a cost-augmented feasible motion planning problem $P = (\hat{\mathcal{Z}}, \mathcal{U}, \hat{\mathbf{z}}_{\text{init}}, \hat{\mathbf{z}}_{\text{goal}}, \hat{\mathcal{Z}}_b, \mathcal{U}_b, \hat{G})$. Here, $\hat{\mathcal{Z}} = \mathcal{Z}_{\text{free}} \times \mathbb{R}^+$, indicating the cost-augmented state space, where $\mathcal{Z}_{\text{free}} \subset \mathcal{Z}$ is a collision-free space. The system state space $\mathcal{Z}_{\text{free}}$ is augmented by an auxiliary cost variable, which measures the accumulated cost from the root \mathbf{z}_{init} , i.e. cost-to-come. \mathcal{U} is the control space, where a control input $\mathbf{u} = [\mathbf{f}_1, \boldsymbol{\tau}_1, \dots, \mathbf{f}_{n_o}, \boldsymbol{\tau}_{n_o}]$ refers to n_o wrenches (force \mathbf{f}_i applied at the i 'th object's CoM and torque $\boldsymbol{\tau}_i$), mimicking an external perturbation in the predictive model. $\hat{\mathbf{z}}_{\text{init}} = (\mathbf{z}_{\text{init}}, 0)$, referring to the augmented root state. The augmented goal set $\hat{\mathcal{Z}}_{\text{goal}} = \mathcal{Z}_\infty \times \mathbb{R}^+$ encompasses an infinitely far away goal set $\mathcal{Z}_\infty \subset \mathcal{Z}$. The augmented state constraint set is denoted by $\hat{\mathcal{Z}}_b = \mathcal{Z}_b \times \mathbb{R}^+$, with unconstrained kinematics $\mathcal{Z}_b = \mathcal{Z}_{\text{free}}$. \mathcal{U}_b is the set of control constraints and $\mathcal{U}_b \subset \mathcal{U}$. \hat{G} refers to the augmented dynamics

TABLE I: Comparison of performance scores w/o positional perturbation.

tasks	# objects	capture score AP		capture score error	success score AP		success score error
		true state	perturbed state		true state	perturbed state	
Pushing	10	1.00	0.96 ± 0.01	-0.04 ± 0.32	0.98	0.97 ± 0.04	-0.04 ± 0.16
Scoping	5	0.96	0.79 ± 0.03	-0.00 ± 0.11	1.00	0.98 ± 0.02	-0.01 ± 0.13

given by

$$\hat{z}' = \begin{bmatrix} z' \\ c' \end{bmatrix} = \begin{bmatrix} G(z, \mathbf{u}) \\ Q(z, \mathbf{u}) \end{bmatrix}. \quad (4)$$

The dynamics is subjected to $z' = G(z, \mathbf{u})$, which is implicitly enforced in Pybullet. $Q(z, \mathbf{u})$ denotes the incremental cost,

$$Q(z, \mathbf{u}) = |W_{\text{ext}}(z, G(z, \mathbf{u}))| \quad (5)$$

$$= |E(G(z, \mathbf{u})) - E(z) - W_{\text{fri}}(z, G(z, \mathbf{u}))|. \quad (6)$$

We represent the incremental cost by the absolute value of the external work $W_{\text{ext}}(z, G(z, \mathbf{u}))$. $W_{\text{fri}}(z, G(z, \mathbf{u}))$ represents the work done by friction when the system transitions from z to $G(z, \mathbf{u})$. Therefore, the cost-to-come c^m of a node z^m is given by

$$c^m = \sum_{i=0}^{N-1} Q(z_i, \mathbf{u}_i), \quad (7)$$

where z_i is the parent node of z_{i+1} , $z_0 = z_{\text{init}}$, $z_N = z^m$.

As detailed in Algo. 1, EST terminates after growing M nodes in the tree (Line 2-3). We thus obtain an energy cost field of M state-cost pairs, $\{(z^0, c^0), \dots, (z^M, c^M)\}$. We thereafter approximate the probabilistic distribution (Line 4-8) and compute the scores $\Omega_{\text{cap}}(z_{\text{init}})$ and $\Omega_{\text{suc}}(z_{\text{init}})$ (Line 9-11) following (1)-(3).

D. Representing High-Dimensional State Space: To effectively manage the high-dimensional state space of multiple objects, $\mathcal{X}^{n_o} \times \dot{\mathcal{X}}^{n_o}$, especially when dealing with large numbers of objects, we utilize a minimum-volume enclosing ellipsoid to represent the objects within the workspace. This approach is inspired by human intuitive strategies, such as how one might handle a cluster of objects with a broom[19], focusing primarily on the centroid and overall spread rather than individual details. This representation is particularly applicable when the objects' scale is significantly smaller than that of the end-effector.

The computational model employs Khachiyani's algorithm to determine the minimum bounding ellipsoid of the centers of mass (CoM) for the n_o objects, denoted as $\{r_x^1, \dots, r_x^{n_o}\}$ [20]. This compact convex set, characterized by ellipsoid radii, central position, and orientation, reduces the dimensionality of the representation. Such a low-dimensional model is predicated on the observation that objects grouped under a common ellipsoidal envelope might tend to behave uniformly under manipulation [1]. Therefore, we hypothesize that different states of objects, $\{x, \dot{x}\} \in \mathcal{X}^{n_o} \times \dot{\mathcal{X}}^{n_o}$, sharing the same ellipsoidal parameters, may demonstrate comparable robustness metrics (energy margin). The efficacy and validity of this ellipsoidal representation and its implications on robustness will be further examined in the subsequent section.

III. EVALUATION

A. Task Description: We have designed two tasks, planar pushing and scooping from a table, to evaluate our proposed method (see Fig. 1-A).

1) Planar Pushing: This task involves the planar manipulation of several objects on a horizontal plane, pushing them towards a wall. Success is defined as achieving a state $z \in \mathcal{Z}_{\text{suc}}$, where $\mathcal{Z}_{\text{suc}} = \{z \in \mathcal{Z} : r_x^i \in \mathcal{R}_{\text{suc}}, \forall i \in \{1, \dots, n_o\}\}$. Here, r_x^i represents the center of mass (CoM) position of the i 'th object. $\mathcal{R}_{\text{suc}} \subset \mathbb{R}^2$ is a designated area near the wall that constitutes the task success region (Fig. 1-B). The capture set, $\mathcal{Z}_{\text{cap}}(z_{\text{init}}) = \{z \in \mathcal{Z} : r_x \in \mathcal{R}_{\text{cap}}(z_{\text{init}}), \forall i \in \{1, \dots, n_o\}\}$, includes states where all objects' CoM positions are within a specific workspace region, $\mathcal{R}_{\text{cap}}(z_{\text{init}})$, shaped like a circular sector swept by the end-effector based on its current instantaneous rotation center.

2) Scooping from a Table: This task involves using a linear scoop gripper to lift several boxes from a table surface. The gripper, powered by a Qbrotic Qbmove actuator[21], features variable stiffness for adaptive gripping. It employs a double parallelogram mechanism on each side to ensure a parallel closing of the fingers without necessitating movement of the robot arm. It uses small RC servos to finely adjust the scoop position for either sliding under objects or achieving precision grasps. Success is defined when all boxes, each with an edge length of 1 cm, are elevated at least 5 cm above the table surface without any boxes slipping or falling from the gripper. A successful capture occurs when boxes remain on the gripper's flat fingertips.

B. Data Generation and Implementation: We collected trajectories and frame-by-frame system states to compute energy margins offline and to compare them with ground-truth labels for robustness and success (Fig. 1-A).

1) Ground-Truth Data Generation: For each task, we generated and recorded 20 trajectories in the simulation by randomizing the initial state of objects and the friction coefficient. From each trajectory, 10 frames ($K = 10$) were selected evenly, and the system states $z_k, k \in \{1, \dots, K\}$ were recorded. A trajectory is labeled as successful (1) if the task objective is achieved in the last recorded frame, i.e., $z_K \in \mathcal{Z}_{\text{suc}}$, otherwise it is labeled as unsuccessful (0). Similarly, a frame state z_k is labeled as captured (1) if it is contained within the capture set $\mathcal{Z}_{\text{cap}}(z_{k'})$ for subsequent frames $k' \in \{k, \dots, k + \hat{k}\}$.

2) Implementation Details: Using Algorithm 1, we computed $\Omega_{\text{cap}}(z_k)$ and $\Omega_{\text{suc}}(z_k)$ for each recorded state z_k across all trajectories. These scores were compared with the ground-truth labels using Average Precision (AP), which indicates performance across classification thresholds. To assess the ability to predict task success, we introduced a trajectory-level score, $\Omega_{\text{suc}}(z, \bar{k})$, calculated as a weighted average of the success scores $\Omega_{\text{suc}}(z_k)$ for the last \bar{k} states,

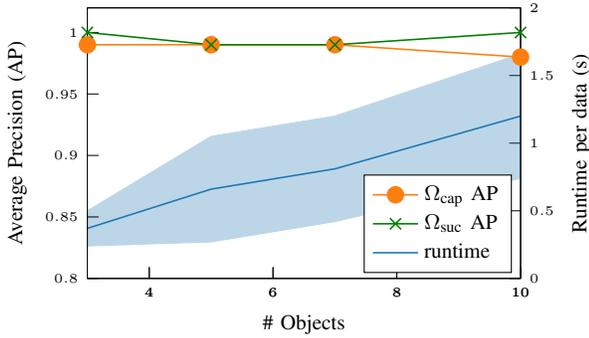


Fig. 2: AP of the performance scores and runtime of the algorithm with respect to the number of objects n_o . Runtime refers to the elapsed time of Algo. 1 given 300 iterations in EST. AP and runtime are computed given the 20 trajectories of the pushing task.

with weights increasing from $k = 1$ to $k = \bar{k}$. This approach resulted in a dataset of 200 state-level data points ($\Omega_{\text{esc}}(z_k), \Omega_{\text{cap}}(z_k)$) and 20 trajectory-level data points $\Omega_{\text{suc}}(z, \bar{k})$ per task, each with corresponding success and capture labels.

C. Robustness against Uncertainty: We conducted an evaluation to assess the robustness of our performance metrics in the presence of object position errors, simulating uncertainties in the perception module. Perturbed positions $\{\hat{r}_x^1, \dots, \hat{r}_x^{n_o}\}$ were resampled on or inside the minimal ellipsoid bounding the objects, while the object velocities were set to 0 in the perturbed state. For each task, we computed the Average Precision (AP) for the performance scores Ω_{cap} and Ω_{suc} given ground-truth object positions (true state) and perturbed positions (perturbed state) (Table I). The score errors were calculated as the differences between the perturbed and true state scores across all data points in the 20 trajectories. The results indicate a strong predictive capability of our metrics with ground-truth object positions, and promising predictions even with perturbed object states. The slightly inferior prediction performance of Ω_{cap} in the scooping task may stem from the empirical definition of the capture set \mathcal{Z}_{cap} . Nevertheless, these results suggest that the ellipsoidal representation has the potential to mitigate the challenges of the curse of dimensionality and perceptual uncertainty in our framework.

D. Ablation Study on Number of Objects: The main results depicted in Fig. 2 reveal two key findings: 1. The runtime exhibits a linear increase with the number of manipulated objects, n_o . 2. Despite the increase in state space dimensionality, our framework maintains strong predictive power. This indicates its scalability and effectiveness in handling scenarios with varying numbers of objects.

IV. CONCLUSION

This workshop contribution has introduced a method for quantifying the robustness of simultaneous multi-object manipulation using energy margins calculated through a kinodynamic motion planning algorithm. While effective, our approach is limited in that it depends on precise system modeling and the computational intensity of dynamic simulations. Moving forward, we aim to apply our robustness characterization in real-world robust manipulation planning

by integrating a heuristic-guided action sampler module into the current framework. Additionally, we are exploring the potential of form-behavior co-design for the scooping gripper, where energy margins are utilized as a behavior quality measure within the design optimization process.

REFERENCES

- [1] Z. Pan, A. Zeng, Y. Li, J. Yu, and K. Hauser, "Algorithms and systems for manipulating multiple objects," *IEEE Trans. Robot.*, vol. 39, no. 1, pp. 2–20, 2022.
- [2] J. Stüber, C. Zito, and R. Stolkin, "Let's push things forward: A survey on robot pushing," *Frontiers in Robotics and AI*, vol. 7, p. 8, 2020.
- [3] N. S. Pollard, "Synthesizing grasps from generalized prototypes," in *Proc. Int. Conf. Robot. Automat.*, vol. 3. IEEE, 1996, pp. 2124–2130.
- [4] C. Ferrari, J. F. Canny, et al., "Planning optimal grasps," in *Proc. Int. Conf. Robot. Automat.*, vol. 3, no. 4, 1992, p. 6.
- [5] H. Tsukamoto, S.-J. Chung, and J.-J. E. Slotine, "Contraction theory for nonlinear stability analysis and learning-based control: A tutorial overview," *Annual Reviews in Control*, vol. 52, pp. 135–169, 2021.
- [6] R. Tedrake et al., "Lqr-trees: Feedback motion planning on sparse randomized trees," in *Robotics: Science and Systems*, vol. 2009, 2009.
- [7] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: Theory and applications," in *European control conference*. IEEE, 2019, pp. 3420–3431.
- [8] M. A. Roa and R. Suárez, "Grasp quality measures: review and performance," *Autonomous robots*, vol. 38, pp. 65–88, 2015.
- [9] Y. Hou and M. T. Mason, "Criteria for maintaining desired contacts for quasi-static systems," in *Proc. Int. Conf. Intell. Robot. Syst.* IEEE, 2019, pp. 6555–6561.
- [10] R. Krug, Y. Bekiroglu, D. Kragic, and M. A. Roa, "Evaluating the quality of non-prehensile balancing grasps," in *Proc. Int. Conf. Robot. Automat.* IEEE, 2018, pp. 4215–4220.
- [11] W. Kuperberg, "Problems on polytopes and convex sets," in *DIMACS Workshop on polytopes*, 1990, pp. 584–589.
- [12] E. Rimon and A. Blake, "Caging 2d bodies by 1-parameter two-fingered gripping systems," in *Proc. Int. Conf. Robot. Automat.*, vol. 2. IEEE, 1996, pp. 1458–1464.
- [13] —, "Caging planar bodies by one-parameter two-fingered gripping systems," *IEEE Int. J. Robot. Res.*, vol. 18, no. 3, pp. 299–318, 1999.
- [14] J. Mahler, F. T. Pokorny, Z. McCarthy, A. F. van der Stappen, and K. Goldberg, "Energy-bounded caging: Formal definition and 2-d energy lower bound algorithm based on weighted alpha shapes," *IEEE Robot. Automat. Lett.*, vol. 1, no. 1, pp. 508–515, 2016.
- [15] Y. Dong and F. T. Pokorny, "Quasi-static soft fixture analysis of rigid and deformable objects," *Proc. Int. Conf. Robot. Automat.*, 2024.
- [16] Y. Dong, X. Cheng, and F. T. Pokorny, "Caging in motion: Characterizing robustness in manipulation through energy margin and dynamic caging analysis," *arXiv preprint arXiv:2404.12115*, 2024.
- [17] S. Makita and W. Wan, "A survey of robotic caging and its applications," *Advanced Robotics*, vol. 31, no. 19-20, pp. 1071–1085, 2017.
- [18] R. R. Ma, W. G. Bircher, and A. M. Dollar, "Modeling and evaluation of robust whole-hand caging manipulation," *IEEE Trans. Robot.*, vol. 35, no. 3, pp. 549–563, 2019.
- [19] S. Elliott and M. Cakmak, "Robotic cleaning through dirt rearrangement planning with learned transition models," in *Proc. Int. Conf. Robot. Automat.* IEEE, 2018, pp. 1623–1630.
- [20] M. J. Todd and E. A. Yildirim, "On khachiyan's algorithm for the computation of minimum-volume enclosing ellipsoids," *Discrete Applied Mathematics*, vol. 155, no. 13, pp. 1731–1744, 2007.
- [21] qbrobotics, "qbmovement advanced," 2024, accessed: 2024-04-20. [Online]. Available: <https://qbrobotics.com/product/qbmovement-advanced/>