

# MiniBEE: A New Form Factor for Compact Bimanual Dexterity

Sharfin Islam<sup>\*†</sup>, Zewen Chen<sup>\*†</sup>, Zhanpeng He<sup>\*‡</sup>, Swapneel Bhatt<sup>†</sup>, Andres Permuy<sup>†</sup>,  
Brock Taylor<sup>†</sup>, James Vickery<sup>†</sup>, Zhengbin Lu<sup>‡</sup>, Cheng Zhang<sup>‡</sup>, Pedro Piacenza<sup>†</sup>, Matei Ciocarlie<sup>†</sup>

[roamlab.github.io/minibee](https://roamlab.github.io/minibee)

**Abstract**—Bimanual robot manipulators can achieve impressive dexterity, but typically rely on two full six- or seven-degree-of-freedom arms so that paired grippers can coordinate effectively. This traditional framework increases system complexity and footprint while only exploiting a fraction of the overall workspace for dexterous interaction. We introduce MiniBEE (Miniature Bimanual End-effector), a compact system in which two reduced-mobility arms (3+ DOF each) are coupled into a kinematic chain that preserves full relative positioning between grippers and enables the entirety of systems workspace to be used for dexterity. To guide our design, we formulate a kinematic dexterity metric to evaluate different kinematic designs. The resulting system supports two complementary modes: (i) wearable kinesthetic data collection with self-tracked gripper poses, and (ii) deployment on a standard robot arm, extending dexterity across its entire workspace. We present kinematic analysis and design optimization methods for maximizing dexterous range, and demonstrate an end-to-end pipeline in which wearable demonstrations train imitation learning policies that perform robust, real-world bimanual manipulation.

**Index Terms**—Bimanual manipulation, dexterous manipulation, wearable imitation learning

## I. INTRODUCTION

In recent years, bimanual robotic manipulators have shown remarkable dexterity. The combination of imitation learning from human demonstrations and two well-articulated kinematic chains has enabled such systems to use simple parallel grippers for highly dexterous tasks [1]–[7], with robustness to initial conditions or perturbations encountered during execution [8]–[10].

To achieve these results, current systems typically rely on the combination of two 6- or 7-degree-of-freedom (DOF) robotic arms. The *dexterous workspace* of the overall system, defined here as the space where the two grippers have complete positioning ability w.r.t. each other, thus lies at the intersection of the two arms’ individual workspaces. As a result, this dexterous workspace is relatively small, and only covers a fraction of each arm’s overall individual workspace. Moreover, this dexterous workspace is fixed relative to the base of the entire system, meaning any object must be brought into this dexterous portion of the system’s workspace to be manipulated. Overall, this traditional framework for

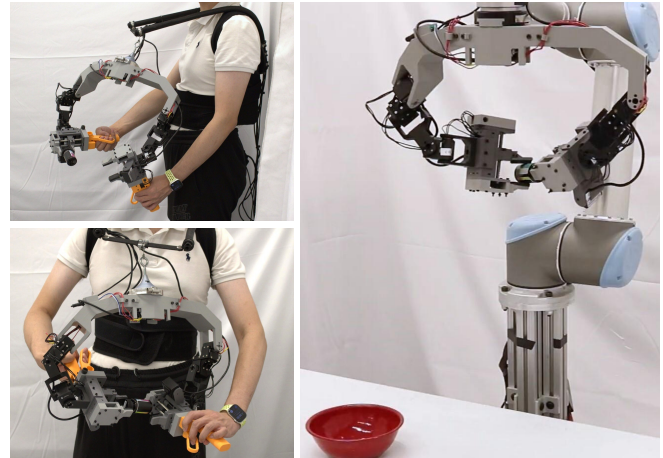


Fig. 1: Overview. The MiniBEE is a bimanual end-effector consisting of two compact arms equipped with grippers. The kinematic chain connecting the two grippers has sufficient degrees of freedom so that, with careful kinematic optimization, the two grippers have full manipulability w.r.t. each other, enabling high dexterity. The MiniBEE is compact enough to be worn and operated for kinesthetic data collection (left) for behavioral cloning, and the trained policies can be deployed with the MiniBEE mounted on a standard robot arm (right), giving the overall system access to almost the entirety of the overall systems workspace for dexterity.

bimanual manipulation is capable of dexterity, but has limitations in terms of efficiency, footprint, and overall mobility.

In this paper, we propose a new approach to bimanual dexterity that aims to address the challenges described above for bimanual systems while preserving many of their benefits. Our fundamental insight is as follows. If each of the arms of a bimanual system has reduced mobility (three+ DOF), the *kinematic chain connecting the two grippers* still has sufficient articulation (6+ DOF) to allow *arbitrary positioning of the grippers w.r.t. each other*, and, with careful kinematic design, we can ensure that the dexterous workspace of the overall system is large enough for bimanual tasks. We dub our system the **MiniBEE**, for Miniature Bimanual End-effector.

Thanks to the reduced articulation of each individual arm, the overall footprint of MiniBEE remains compact and lightweight. In turn, this provides two key benefits:

- The MiniBEE can be worn directly by an operator, allowing for easy and mobile kinesthetic data collection. When this data is used to train autonomous policies, the encoders in the kinematic chain provide full information

<sup>\*</sup> joint first authorship

<sup>†</sup>Dept. of Mechanical Engineering, <sup>‡</sup>Dept. of Computer Science  
Columbia University, New York, NY 10027, USA

Corresponding email: [si2395@columbia.edu](mailto:si2395@columbia.edu)

This work was supported in part by the NSF under awards PFI-232975 and CMMI-2037101.

on the relative pose of the two grippers, removing the need for external tracking or SLAM while the MiniBEE is manipulating an object held in or transferred between its grippers.

- The MiniBEE can also be mounted on a commodity robot arm, for example when a trained policy is deployed. This means that the entire reachable area of the robot arm instantly becomes available for the MiniBEE, and the dexterous workspace (i.e. the area where the two grippers have complete positioning authority w.r.t. each other) is scaled by the overall reach of the robot arm.

The MiniBEE functions both as wearable mechanism for easy data collection with self-tracking dexterity, and as a bimanual end-effector allowing a single traditional robot arm to become dexterous over the entirety of its workspace. Overall, our main contributions are as follows.

- To the best of our knowledge, this is the first example of reduced mobility arms enabling full relative mobility for bimanual tasks in a system that is compact enough for wearable collection of kinesthetic demonstrations.
- We introduce a kinematic analysis tool that allows a designer to maximize the dexterous workspace while limiting the total number of DOF, and use these to compare multiple possible design variants.
- We demonstrate the end-to-end pipeline consisting of wearable kinesthetic data collection followed by training standard imitation learning policies which leverage the proprioceptive component of the demonstrations to achieve robust, dexterous bimanual manipulation with large reach in the real world.

## II. RELATED WORK

Bimanual systems offer several advantages for high dexterity, such as accessible off-the-shelf hardware and reliable joint control. With the appropriate kinematic structure and control policy, such systems can perform a wide range of dexterous robotic tasks. A central challenge, however, lies in teaching robots such coordination. With recent advances in learning-based control, success in bimanual manipulation increasingly depends on building pipelines that can collect large, high-quality datasets. Most existing approaches to data collection fall into two categories: **teleoperated** [4], [5], [11]–[14] and **in-the-wild** [8], [15]–[18].

Teleoperation enables high-fidelity demonstrations but requires expensive setup and operator training before consistent demonstrations can be produced [4], [5], [19]. For example, ALOHA introduces a low-cost bi-manual leader and follower system [11] and demonstrates robust bi-manual skills. The leader is composed of un-actuated linkages of the same kinematics as follower robot and is controlled by a human operator. Once available, data collected from such systems typically transfers well, since it is collected directly from the target robot. However, such systems typically require a large-footprint leader-follower setup deployed at the data collection site. Moreover, teleoperated follower robots takes a significant amount of practice to get accustomed to the lack of proprioceptive feedback for the human operator [20].

By contrast, in-the-wild methods aim to bypass these barriers by collecting data on disembodied end-effectors. The DobbE system [17] performed data collection with a single parallel-gripper that closely matched the real robot’s embodiment. For bi-manual manipulation, the UMI system [7] introduced a handheld paradigm with two parallel grippers and cameras, allowing users to record bi-manual demonstrations. Similarly, DeXwild [18] collected data in-the-wild, but using motion capture gloves and deploying on anthropomorphic robot hands. However, since no robot embodiment is present during collection, significant effort is still required to map these demonstrations onto real robots. Furthermore, demonstrations that prove unfeasible under the kinematic constraints of the complete robot must be discarded.

Recently, a third category has emerged: wearable bi-manual data collection systems [6], [21], [22]. These systems aim to collect robot-relevant data without requiring direct access to a robot, instead using wearable hardware that mirrors the robot’s embodiment. By doing so, they reduce embodiment mismatch while retaining the scalability of human demonstrations. This combines the strengths of both prior paradigms: scalable data collection with minimal post-processing, and demonstrations that align closely with robot embodiment.

Examples of dual arm systems within this wearable category include AirExo [6], a 14-DOF wearable exoskeleton for bi-manual data collection, and the dual-arm teleoperation system in [22]. Both consist of long kinematic chains that are fastened to the torso and are embedded with joint encoders. These systems designed to provide mobile demonstrations for traditional dual-arm robots, while mitigating the footprint of the system as much as possible. However, they also follow the framework of two highly articulated arms fastened to a fixed world frame, which increases complexity and footprint, while only being able to leverage a small portion of the overall systems workspace for dexterity. In this work, we introduce a new wearable system within this category, with a novel kinematic design, that is designed to take full advantage of the system’s workspace. By reducing redundancy between the grippers, we achieve a significantly more compact and lightweight design, which makes the manipulator both wearable for data collection and mountable on a robot arm for large-reach policy execution.

## III. DESIGN AND OPTIMIZATION

The core idea of the MiniBEE is to realize a compact bimanual manipulator in which each arm individually has reduced mobility, but the combined kinematic chain formed between the two grippers retains sufficient articulation (6+ DOF) to achieve extensive relative mobility. In other words, through careful intrinsic design, the MiniBEE provides a dexterous workspace in which one gripper can be positioned and oriented with respect to the other with high flexibility. Absolute positioning in the environment can then be provided by mounting the system on a standard robotic arm, which

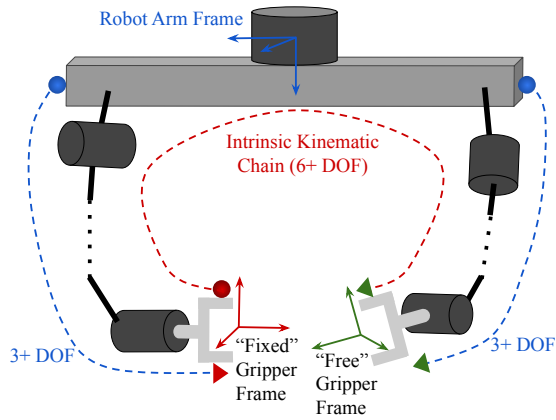


Fig. 2: MiniBEE concept. We use a bimanual system designed to mount on a larger robot arm. When designing the MiniBEE, we rely on a kinematic convention which treats one gripper as the base coordinate frame (the “fixed” gripper) and the other as the tip of the chain (the “free” gripper). Rather than computing each gripper’s pose relative to a robot tool frame (blue), we can then solve kinematic queries directly between the two grippers, relying on the fact that the complete kinematic chain connecting them (red) has sufficient DOFs for relative dexterity.

then expands the kinematic chain for each gripper relative to the overall world frame. This concept is illustrated in Fig. 2.

While 6+ intrinsic DOF are the theoretical minimum for arbitrary relative positioning, we find that the *size* of the resulting dexterous workspace is strongly influenced by additional kinematic design choices, such as link lengths and joint limits, that will influence self-collisions and overall kinematic mobility at desired poses. In this section, we describe our methodology for expanding this workspace while maintaining a low intrinsic DOF count, thereby keeping the MiniBEE compact and lightweight. In the following section, we show how these design properties enable efficient data collection and manipulation via imitation learning.

### A. Kinematic Analysis

Conventional bimanual robots are described as two independent chains rooted at a shared base. Here, however, we focus exclusively on the *intrinsic* DOF of the MiniBEE and their ability to control the relative pose of the two grippers. To this end, we adopt a different convention: the system is modeled as a single kinematic chain, with the base frame coincident with one gripper and the tool frame at the other gripper (Fig. 2).

To evaluate relative dexterity, we introduce the **Kinematic Dexterity (KD) metric**. Conceptually, the KD metric is designed to test the ability of one gripper to achieve a wide range of useful poses w.r.t. the other gripper.

The KD metric is defined as follows. One gripper, referred to as the “fixed” gripper, is treated as stationary and establishes the origin of the coordinate frame. The other gripper, referred to as the “free” gripper, is evaluated based on its ability to attain a diverse set of poses relative to this fixed frame. To this end, we uniformly sample candidate target poses, consisting of both translational and rotational compo-

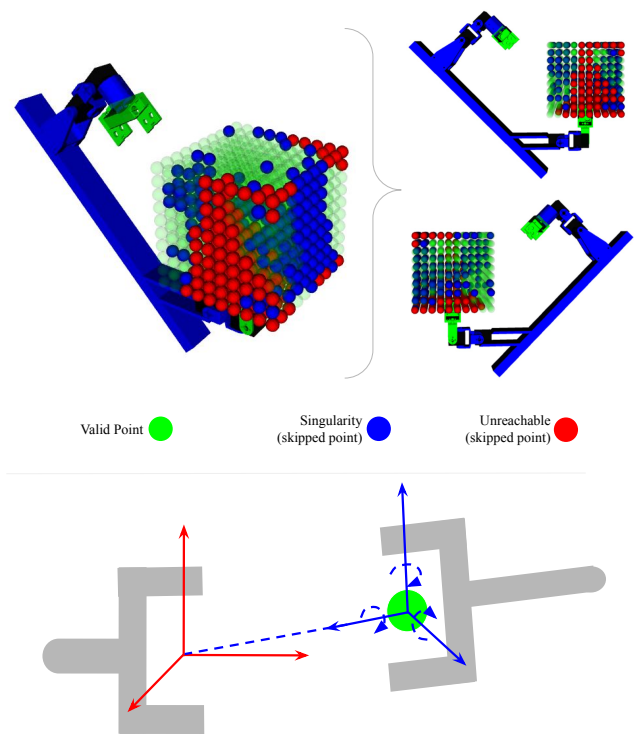


Fig. 3: We evaluate the kinematic dexterity between two grippers of a bi-manual kinematic chain. One of the grippers in the chain, dubbed the fixed gripper, serves as the base of our coordinate frame. We then set a desired volume for a workspace around the fixed gripper uniformly sample a set of points in this volume. We use each of these points as a desired position for the “free” gripper, oriented such that it points towards the fixed gripper (bottom image). If an IK solution for this query exists, obeys joint limits and self-collisions, and does not place the robot near a singularity, we consider the query point as reachable. The ratio of reachable to non-reachable points in the workspace is used as a metric for the relative kinematic dexterity between two grippers that are connected by a single kinematic chain.

nents, within a predefined rectangular region centered on the fixed gripper. For each sampled pose, we employ standard inverse kinematics (IK) to assess feasibility, considering both joint-limit constraints and potential self-collisions. The KD metric is then defined as the proportion of sampled poses that are kinematically feasible under these conditions. The procedure is described in Alg. 1 and also illustrated in Fig. 3

We note that other formulations of dexterity metrics exist (e.g., manipulability ellipsoids, ergonomics-inspired indices [23]–[25]); however, we are interested here specifically in the relative positioning and orienting ability most relevant to bimanual tasks. For collision- and joint limit-aware IK, we use the KDL plugin [26] in the MoveIt library [27], which ensures realistic consideration of such constraints.

### B. Design Characterization

With the KD metric established, we evaluate multiple possible MiniBEE configurations against each other, and also against an existing bimanual platform (Fig. 4). We start with a configuration with 6 intrinsic DOFs, the theoretical

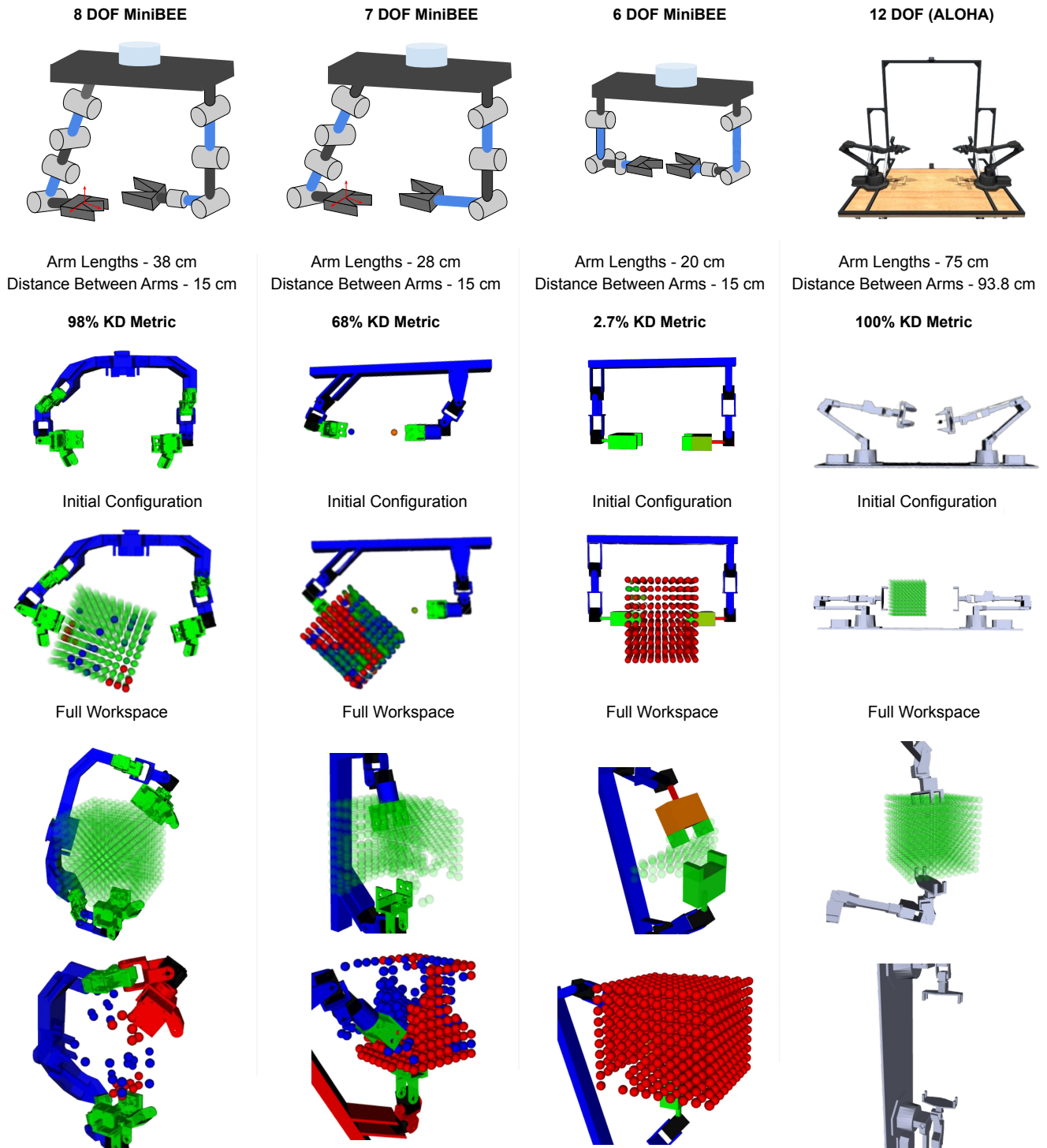


Fig. 4: Kinematic analysis for possible MiniBEE configurations as well as a state of the art bimanual system. For each configuration, we follow the procedure outlined in Algorithm 1 to compute the KD-metric for a desired fixed rectangular workspace with a side length of 20 cm. Successful desired points in the workspace are shown in green, points skipped due to IK solutions being near singularities are shown in blue, and points where there are no collision-free IK solutions are shown in red. The 8-DOF MiniBEE performed the best, as expected, and was comparable in performance to traditional bimanual systems that use two full mobility arms. Reducing the configuration to 7 and 6-DOF reduced the dexterity score significantly. For all experiments shown in the paper, we used the 8-DOF configuration, as it provided the best balance between complexity, size, and performance. The state of the art robot performs perfectly on our metric, but has 6 degrees of redundancy between the grippers. Overall, our KD metric correlates with the total degrees of freedom, and overall kinematic redundancy between the two gripper frames.

---

**Algorithm 1** Computation of the KD metric

---

**Input:** Robot description (kinematics and joint limits), target workspace size and resolution, singularity threshold  $\epsilon$

**Output:** Kinematic Dexterity metric  $KD \in [0, 1]$

- 1: Create a grid of points by sampling the target workspace at the desired resolution (see Fig. 3)
  - 2:  $N_{\text{total}} \leftarrow$  total number of points in grid
  - 3:  $N_{\text{valid}} \leftarrow 0$
  - 4: **for** each point  $\mathbf{p}_i$  in the grid **do**
  - 5:   Compute free gripper pose  $\mathbf{P}_i(\mathbf{p}_i)$  using Alg. 2
  - 6:   Robot pose  $\mathbf{q}_i = \text{InverseKinematics}(\mathbf{P}_i)$
  - 7:   **if** no  $\mathbf{q}_i$  exists **then**
  - 8:     Skip this point (point is not reachable)
  - 9:    $\mathbf{J}_i = \text{Jacobian}(\mathbf{q}_i)$
  - 10:   Remove  $\mathbf{J}_i$  columns corresp. to joints near limits
  - 11:   Compute singular values of  $\mathbf{J}_i$
  - 12:   **if** smallest singular value  $< \epsilon$  **then**
  - 13:     Skip this point (robot is near a singularity)
  - 14:   Count this point as valid:  $N_{\text{valid}} \leftarrow N_{\text{valid}} + 1$
  - 15:  $KD \leftarrow N_{\text{valid}} / N_{\text{total}}$
- 

---

**Algorithm 2** Computation of free gripper poses

---

**Input:** Workspace point  $\mathbf{p}_i$

**Output:** Corresponding free gripper pose  $\mathbf{P}_i$

- 1: Compute rotation  $\mathbf{R}(\mathbf{p}_i)$  such that approach direction (z-axis) of free gripper points towards origin
  - 2:  $\mathbf{P}_i \leftarrow (\mathbf{R}(\mathbf{p}_i), \mathbf{p}_i)$
- 

minimum sufficient for relative dexterity. However, in the presence of realistic constraints such as self-collisions and joint limits, we find that adding more intrinsic DOFs progressively increases redundancy between the grippers and thus relative dexterity. We thus also test 7- and 8-DOF configurations of the MiniBEE using the same metric.

For each configuration, we used realistic joint limits and collision geometries based on available off-the-shelf servomotors (in our case, the Dynamixel X family). Of course, given a number of DOFs, many kinematic configurations are possible. Here, we selected the exact kinematics in each case empirically, based on design intuition. The KD metric allowed to quickly evaluate candidates, and select the best one in each class. A very compelling direction for future work would be to automate this process, and perform a full-fledged design optimization process with our KD metric as a fitness function.

While not directly comparable to the MiniBEE, we also compute the KD metric for an existing bimanual system, namely the ALOHA manipulator [5]. We note that the ALOHA follower robot is not intended for wearability and kinesthetic teaching. However, with 12 total DOFs between the two arms, it represents a standard of performance in terms of relative dexterity. As such, it provides us with useful grounding information for the capabilities of the MiniBEE.

For each system, we applied Algorithm 1 with a fixed

20 cm cubic workspace, defined with one face centered at the fixed gripper and extending toward the free gripper. The same workspace size was used for all MiniBEE variants and the ALOHA system, despite the latter’s longer link lengths, to ensure a fair comparison.

Across MiniBEE variants, as expected, we find that more intrinsic DOF yields higher KD scores. The 8-DOF configuration maintained feasible kinematics over most of the workspace, with singularities appearing only when some of the joints become unusable to the robot as they hit their limits. Reducing to 7 or 6 DOFs sharply lowered performance: the 7-DOF system encountered frequent singularities and self-collisions, while the 6-DOF system often failed to find any solutions for many of the points. These results show that some redundancy between grippers is essential for avoiding self-collision, limiting singularities, and overall preserving dexterity. We therefore select the 8-DOF MiniBEE as it offers the best trade-off between compactness, complexity, and performance, while remaining competitive with more traditional bi-manual systems.

The ALOHA system provides an interesting reference point. Its 12-DOF structure attains a perfect KD score. Although the 8-DOF MiniBEE achieves a slightly lower score, it is still comparable with fewer joints and a much smaller footprint overall. Based on this analysis, we believe that the MiniBEE provides relative dexterity comparable to state of the art system, while remaining compact and lightweight enough to be either worn for kinesthetic teaching or mounted on a robot arm for high reach, features that we put to use in the next section.

In addition to kinematic capabilities, static payload capacity is an important mechanical characteristic. For our 8-DOF design used in the experiments (Fig. 5), we calculated payload capacity for both the worst-case and a natural manipulation pose. We assumed that all motors operate at 30% of stall torque given by the manufacturer, link masses are concentrated at the joints, and bracket/fastener masses are negligible. The worst-case pose yields a payload of 18g, including the 460g gripper mass, though this pose is unlikely during manipulation. In a more natural pose, with grippers facing each other and the UR5 tool frame mostly parallel to the floor, the payload increases to 168g in addition to the gripper mass. This is sufficient for our light-object, fine-manipulation tasks, though future iterations could improve capacity with lighter grippers or consideration of payload in our design analysis.

#### IV. ACHIEVING DEXTEROUS MANIPULATION

Having described the overall design of the MiniBEE, we now turn to the goal of achieving dexterous bimanual manipulation. In particular, we rely on imitation learning from human demonstrations, a method recently shown to achieve skilled manipulation in this context [4]–[7].

The fundamental advantage of MiniBEE, as described in the previous section, is that it enables intrinsic bimanual manipulation with fewer DOF than existing alternatives. It is compact, lightweight, and can be worn directly by an

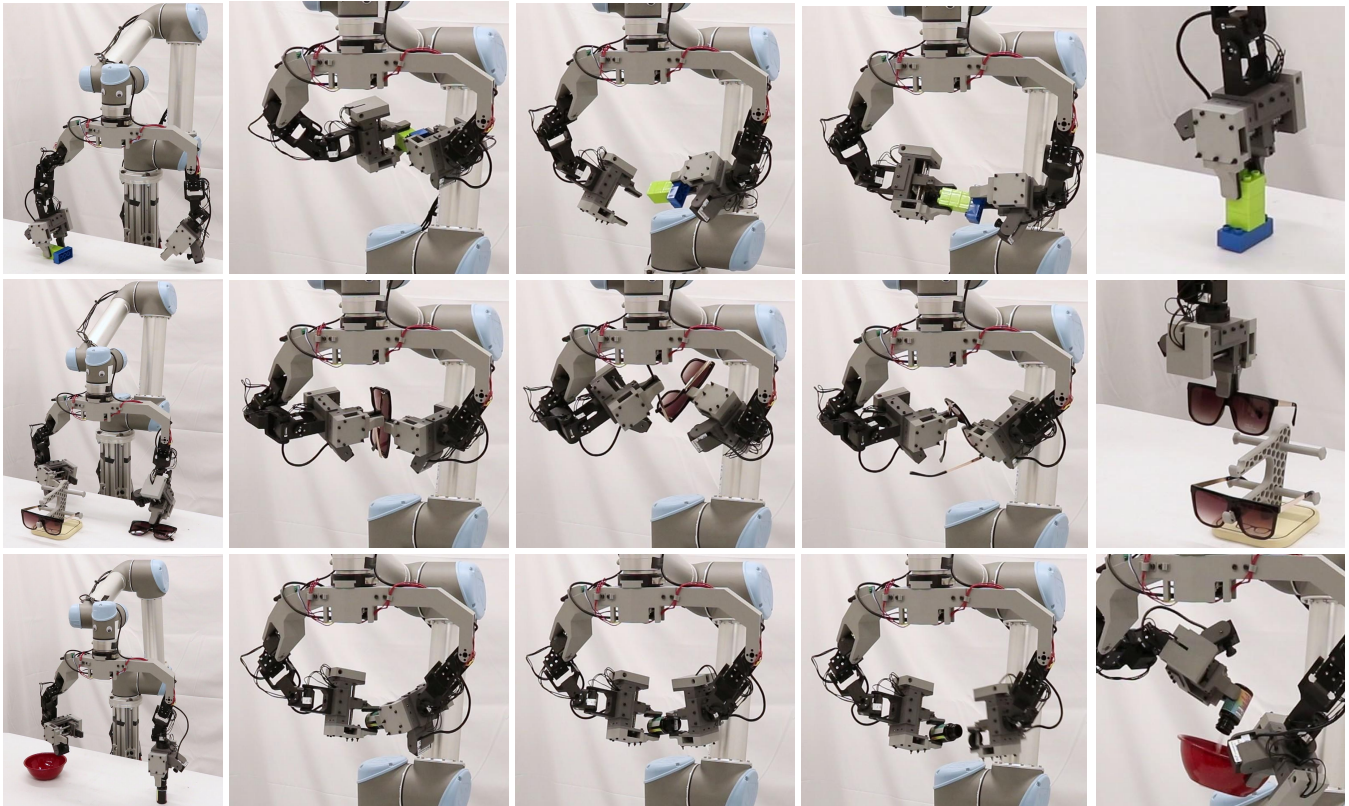


Fig. 5: We conduct a set of three manipulation experiments with MiniBEE. The first task, shown in the first row, is to grasp an object from the table, reorient it by handing it off to the other gripper, and place it on the table in a different pose. The second task, shown in the second row, is for MiniBEE to pick up a pair of folded sunglasses from the table, unfold them with the other gripper, and then place on a display rack. The third task is to pick up a closed pill bottle, unscrew the cap, pick up a bowl with the other gripper, and then empty the contents of the bottle into the bowl. MiniBEE performs these task with a high rate of success, and overall this set of experiments highlights the ability for our design to achieve bi-manual dexterity.

operator for kinesthetic data collection. This significantly reduces the difficulty of collecting demonstrations compared to systems with large robot arms, and even more so compared to leader-follower systems requiring four arms in total.

Compared to "in-the-wild" systems where the operator uses robot end-effectors, MiniBEE offers several advantages. First, the relative pose of the grippers is tracked through encoders in the kinematic chain, eliminating the need for external tracking or SLAM during intrinsic manipulation. Second, demonstrations are collected under the full set of kinematic constraints of the real system, ensuring they are feasible on hardware. Finally, MiniBEE is lightweight enough to fit within the payload of standard robot arms, extending reach for larger workspaces, unlike other wearable exoskeletons such as [6].

The self-tracking capabilities of MiniBEE apply during intrinsic manipulation, where the object is held by, or transferred between, the grippers. When manipulating an object fixed in an external reference frame, intrinsic encoders alone are insufficient. If one gripper remains fixed, intrinsic DOFs still suffice, but if both must move w.r.t. the external frame, additional tracking is needed, similar to in-the-wild systems. For this paper, our policies will focus on demonstrating the intrinsic dexterity between the two grippers, not the dexterity

relative to the environment.

In this section, we describe our setup for data collection and policy training, and illustrate MiniBEE's performance on real-world dexterous bimanual tasks.

#### A. Data Collection and Policy Training

To collect data for policy learning, we leverage the fact that the same device can be used for data collection and policy deployment. The MiniBEE is worn and directly controlled by a human operator. A custom back brace (Fig. 1) supports the weight of the MiniBEE and stabilizes the system during demonstrations. Two ergonomic handles allow the operator to control the arms, and integrated triggers provide continuous control of the grippers, enabling natural modulation of grasp force and aperture throughout each task. The parallel grippers are equipped with RGB cameras, and the joints are driven by servos with embedded encoders to record joint angles. The total weight of MiniBEE, with the handles attached to the grippers, is 2.0kg and can be easily supported by our custom back brace, shown in Fig 1, and also the UR5 robot arm.

As mentioned earlier, in this paper, we focus on the intrinsic dexterity between the two grippers. Therefore, we do not collect information about the global positioning of the MiniBEE relative to an external fixed frame in the

environment. As such, the components of a task where an object is first acquired from the environment, or placed back after task completion, are executed open loop. However, any bimanual component of the task is autonomous, using a policy based on the data collection procedure described here.

Relative gripper poses are continuously self-tracked during demonstration collection, and all sensor signals are logged throughout the trajectory at 40 Hz. The observation space consists of RGB images from two cameras mounted on the grippers, as shown in Fig. 1. The images are streamed at 640×480, then cropped to 480×360 to focus on the gripper and manipulated objects, and then downsampled to 320×240. The joint angles are recorded from the encoders embedded in the servos. The observation space also includes the relative pose between the two grippers during policy training and deployment. The action space comprises binary commands for gripper aperture and continuous commands for arm joint positions, directly derived from the operator’s handle inputs. This representation enables the policy to reproduce smooth, continuous manipulation behaviors rather than discrete open/close events.

To address the distribution shift between training and deployment—where handles are visible in demonstration observations but absent during robot execution—we additionally collect a small set of replayed trajectories, which are generated by the recorded data at 20 Hz. These provide handle-free visual inputs while preserving the underlying motion, helping the learned policy adapt to the deployment setting.

For policy training, we use Diffusion Policy (DP) [8], which models the trajectory distribution with a conditional denoising diffusion process. In our implementation, we use a longer observation horizon for low-dimensional state inputs (e.g., gripper poses and joint angles) to better capture temporal dependencies, while restricting image inputs to the most recent frame for computational efficiency. The Diffusion Policy works at 5 Hz, which is trained by the down-sampled data.

### B. Experiments and Performance

We used the method above on a total of 3 tasks, each of them requiring bimanual dexterity for completion. All tasks are illustrated in Fig. 5. For each task, there is a component where a single gripper must pick the object from a table in the environment. As stated, we are focusing only on demonstrating the intrinsic dexterity of our system, and therefore any pick up or placing of objects from the environment is a recorded trajectory. However, any manipulation between the grippers is done by a trained diffusion policy.

1) *Task 1*: Object re-orientation via double handoff. Here, the MiniBEE re-orientes a grasped object by first transferring it to the other gripper, then transferring it back to the original gripper in a different pose. Such re-orientation is commonly needed in many applications, where the grasp originally available for an object is not the same one needed for later use.

2) *Task 2*: Unfolding sunglasses. Here, the MiniBEE holds a folded pair of sunglasses in one gripper, then uses the other gripper to unfold the glasses, followed by placing the unfolded glasses on a shelf. This task highlights non-prehensile relative dexterity.

3) *Task 3*: Unscrewing and emptying bottle. Here, the MiniBEE uses both gripper to first unscrew a bottle, then to pour its contents into a bowl. This task highlights both prehensile and non-prehensile relative dexterity,

For each of these tasks, we collect a total of  $N = 45$  demonstrations via the wearable MiniBEE as described in the previous section. We then augment the training set with  $M = 5$  replay trajectories. We use these demonstrations to train autonomous policies for all the bimanual components of the tasks. Finally, we mount the MiniBEE on a robot arm, and test autonomous execution of the complete tasks.

Our results show a 20/20 success rate for Task 1, 19/20 success for Task 2, and 18/20 success rate for Task 3. For Task 2, the lone failure was due to an imprecise grasp of the temples (or folding arms) of the glasses. For Task 3, both failures were due to incomplete lid unscrewing. Representative executions of all tasks are shown in Fig. 5 as well as the Supplementary Video. Overall, these examples showcase the ability of the MiniBEE system to achieve dexterous bimanual manipulation based on demonstrations collected kinesthetically via a wearable system, leveraging its compact and lightweight nature.

## V. CONCLUSION AND FUTURE WORK

In this paper we have introduced MiniBEE, a concept for a new form factor for bimanual manipulation. MiniBEE attempts to combine the best features of both in-the-wild and leader-follower systems for data collection of bimanual manipulation. Like leader-follower systems, it comprises a two-gripper kinematic chain allowing fully self-tracked relative poses of the two grippers. However, through careful kinematic design analysis, MiniBEE uses a reduced articulation model, while still maintaining a level of dexterity (computed based on the relative positioning ability of the two grippers) comparable to much larger systems that use many more articulations.

This kinematic approach in turn enables a compact and lightweight design of the overall MiniBEE, which can be either worn directly by an operator (for kinesthetic data collection) or mounted on a robot arm (for high reach during autonomous policy execution). We demonstrate both of these modes by collecting kinesthetic data and training policies for multiple tasks requiring bimanual dexterity, both prehensile and non-prehensile. Our results show that the MiniBEE is indeed kinematically capable of such tasks, executing the policies trained from wearable data collection with high success rates.

While MiniBEE exhibits a high level of bimanual dexterity even in a compact package, additional areas of improvement remain. In particular, in future work we aim to implement self-tracking of uni-manual tasks executed relative to a fixed frame in the environment (possible in the case where the

other gripper is kept fixed relative to the manipulation frame) in order to allow closed-loop policies for object acquisition from the environment. Other possible areas of improvement include better models of the MiniBEE position control system, to compensate for different gravity effects during data collection (grippers held by the operator) and policy execution (free-floating grippers). Finally, our KD metric used to assess different kinematics can be improved to include notions of torque, payload, and mass. Nevertheless, we believe MiniBEE is a novel approach to bimanual dexterity and a step towards highly dexterous bimanual systems that can easily be deployed for a wide range of applications in human environments.

## REFERENCES

- [1] Y. Hou, K. Sahari, and D. How, "A review on modeling of flexible deformable object for dexterous robotic manipulation," *International Journal of Advanced Robotic Systems*, vol. 16, no. 3, 2019.
- [2] D. Seita, S. Srinivasa, D. Fox, and S. Levine, "Speedfolding: Learning efficient bimanual folding of garments," in *Proceedings of the 2022 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 5581–5588.
- [3] W. Chen, J. Liang, and D. Sun, "Reinforcement learning of bimanual manipulation strategies," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 4368–4375. [Online]. Available: <https://ieeexplore.ieee.org/document/8967763>
- [4] Z. Fu, T. Z. Zhao, and C. Finn, "Mobile aloha: Learning bimanual mobile manipulation with low-cost whole-body teleoperation," in *Conference on Robot Learning (CoRL)*, 2024.
- [5] T. Z. Zhao, S. Schmidgall, J. W. Kim, A. Deguet, M. Kobilarov, A. Krieger, and C. Finn, "Aloha 2: An enhanced low-cost hardware for bimanual teleoperation," *arXiv preprint arXiv:2405.02292*, 2024.
- [6] H. Fang, H.-S. Fang, Y. Wang, J. Ren, J. Chen, R. Zhang, W. Wang, and C. Lu, "Airexo: Low-cost exoskeletons for learning whole-arm manipulation in the wild," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 15 031–15 038.
- [7] C. Chi, Z. Xu, C. Pan, E. Cousineau, B. Burchfiel, S. Feng, R. Tedrake, and S. Song, "Universal manipulation interface: In-the-wild robot teaching without in-the-wild robots," 2024. [Online]. Available: <https://arxiv.org/abs/2402.10329>
- [8] C. Chi, Z. Xu, S. Feng, E. Cousineau, Y. Du, B. Burchfiel, R. Tedrake, and S. Song, "Diffusion policy: Visuomotor policy learning via action diffusion," 2024. [Online]. Available: <https://arxiv.org/abs/2303.04137>
- [9] C. Zhu, R. Yu, S. Feng, B. Burchfiel, P. Shah, and A. Gupta, "Unified world models: Coupling video and action diffusion for pretraining on large robotic datasets," 2025. [Online]. Available: <https://arxiv.org/abs/2504.02792>
- [10] T. L. Team, "A careful examination of large behavior models for multitask dexterous manipulation," 2025. [Online]. Available: <https://arxiv.org/abs/2507.05331>
- [11] A. . Team, J. Aldaco, T. Armstrong, R. Baruch, J. Bingham, S. Chan, K. Draper, D. Dwivedi, C. Finn, P. Florence, S. Goodrich, W. Gramlich, T. Hage, A. Herzog, J. Hoech, T. Nguyen, I. Storz, B. Tabanpour, L. Takayama, J. Tompson, A. Wahid, T. Wahrburg, S. Xu, S. Yaroshenko, K. Zakka, and T. Z. Zhao, "Aloha 2: An enhanced low-cost hardware for bimanual teleoperation," 2024. [Online]. Available: <https://arxiv.org/abs/2405.02292>
- [12] J. J. Liu, Y. Li, K. Shaw, T. Tao, R. Salakhutdinov, and D. Pathak, "Factr: Force-attending curriculum training for contact-rich policy learning," 2025. [Online]. Available: <https://arxiv.org/abs/2502.17432>
- [13] S. Dass, K. Pertsch, H. Zhang, Y. Lee, J. J. Lim, and S. Nikolaidis, "Pato: Policy assisted teleoperation for scalable robot data collection," 2023. [Online]. Available: <https://arxiv.org/abs/2212.04708>
- [14] Y. Park, J. S. Bhatia, L. Ankile, and P. Agrawal, "Dexhub and dart: Towards internet scale robot data collection," 2024. [Online]. Available: <https://arxiv.org/abs/2411.02214>
- [15] Zhaxizhuoma, K. Liu, C. Guan, Z. Jia, Z. Wu, X. Liu, T. Wang, S. Liang, P. Chen, P. Zhang, H. Song, D. Qu, D. Wang, Z. Wang, N. Cao, Y. Ding, B. Zhao, and X. Li, "Fastumi: A scalable and hardware-independent universal manipulation interface with dataset," 2025. [Online]. Available: <https://arxiv.org/abs/2409.19499>
- [16] M. Xu, H. Zhang, Y. Hou, Z. Xu, L. Fan, M. Veloso, and S. Song, "Dexumi: Using human hand as the universal manipulation interface for dexterous manipulation," 2025. [Online]. Available: <https://arxiv.org/abs/2505.21864>
- [17] N. M. M. Shafiqullah, A. Rai, H. Etukuru, Y. Liu, I. Misra, S. Chintala, and L. Pinto, "On bringing robots home," *arXiv preprint arXiv:2311.16098*, 2023.
- [18] T. Tao, M. K. Srirama, J. J. Liu, K. Shaw, and D. Pathak, "Dexwild: Dexterous human interactions for in-the-wild robot policies," *Robotics: Science and Systems (RSS)*, 2025.
- [19] M. Connan, B. Henze, A. Gigli, C. Castellini, and M. A. Roa, "A wearable, ultralight interface for bimanual teleoperation of a compliant, whole-body-controlled humanoid robot," 2018. [Online]. Available: <https://api.semanticscholar.org/CorpusID:207811274>
- [20] N. T. Tanacar, M. H. Mughrabi, A. U. Batmaz, D. Leonardis, and M. Sarac, "The impact of haptic feedback during sudden, rapid virtual interactions," in *2023 IEEE World Haptics Conference (WHC)*, 2023, pp. 64–70.
- [21] H.-S. Fang, B. Romero, Y. Xie, A. Hu, B.-R. Huang, J. Alvarez, M. Kim, G. Margolis, K. Anbarasu, M. Tomizuka, E. Adelson, and P. Agrawal, "Dexop: A device for robotic transfer of dexterous human manipulation," *arXiv preprint arXiv:2509.04441*, 2025.
- [22] F. Falck, K. Larppichet, and P. Kormushev, "De vito: A dual-arm, high degree-of-freedom, lightweight, inexpensive, passive upper-limb exoskeleton for robot teleoperation," in *Towards Autonomous Robotic Systems*, K. Althoefer, J. Konstantinova, and K. Zhang, Eds. Cham: Springer International Publishing, 2019, pp. 78–89.
- [23] J. J. Yang, W. Yu, J. Kim, and K. Abdel-Malek, "On the placement of open-loop robotic manipulators for reachability," *Mechanism and Machine Theory*, vol. 44, no. 4, pp. 671–684, 2009. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0094114X08001006>
- [24] M. Amani and R. Akhavian, "Intelligent ergonomic optimization in bimanual worker-robot interaction: A reinforcement learning approach," *Automation in Construction*, vol. 168, p. 105741, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0926580524004771>
- [25] S. Selha, P. Dupont, R. D. Howe, and D. F. Torchiana, "Dexterity optimization by port placement in robot-assisted minimally invasive surgery," in *Telemanipulator and Telepresence Technologies VIII*, M. R. Stein, Ed., vol. 4570, International Society for Optics and Photonics. SPIE, 2002, pp. 97 – 104. [Online]. Available: <https://doi.org/10.1117/12.454734>
- [26] R. Smits, H. Bruyninckx, *et al.*, "Kinematics and dynamics library (kdl)," <http://www.orocos.org/kdl>, 2001–, part of the Orococos Project.
- [27] D. Coleman, I. Sukan, S. Chitta, and N. Correll, "Reducing the barrier to entry of complex robotic software: a moveit! case study," 2014. [Online]. Available: <https://arxiv.org/abs/1404.3785>