ROBOTIC MANIPULATION

Complex manipulation with a simple robotic hand through contact breaking and caging

Walter G. Bircher*, Andrew S. Morgan, Aaron M. Dollar

Humans use all surfaces of the hand for contact-rich manipulation. Robot hands, in contrast, typically use only the fingertips, which can limit dexterity. In this work, we leveraged a potential energy–based whole-hand manipulation model, which does not depend on contact wrench modeling like traditional approaches, to design a robotic manipulator. Inspired by robotic caging grasps and the high levels of dexterity observed in human manipulation, a metric was developed and used in conjunction with the manipulation model to design a two-fingered dexterous hand, the Model W. This was accomplished by simulating all planar finger topologies composed of open kinematic chains of up to three serial revolute and prismatic joints, forming symmetric two-fingered hands, and evaluating their performance according to the metric. We present the best design, an unconventional robot hand capable of performing continuous object reorientation, as well as repeatedly alternating between power and pinch grasps—two contact-rich skills that have often eluded robotic hands—and we experimentally characterize the hand's manipulation capability. This hand realizes manipulation motions reminiscent of thumb–index finger manipulative movement in humans, and its topology provides the foundation for a general-purpose dexterous robot hand.

INTRODUCTION

The dexterity of a robotic hand can be greatly increased when all of its surfaces are used for manipulation, rather than just the fingertips (*1*). Despite this, most manipulation research is rooted in the assumption of fixed contact points between the tips of the fingers and the object, a precedent established many decades ago (*2*, *3*). This is in part due to the complexity involved in modeling and controlling nonfixed contacts, which would be necessary to prevent object ejection during force closure–based rolling or sliding fingertip manipulation. Accurately predicting the motion of sliding contacts depends on knowledge of both force magnitude and direction at the point of contact, in addition to the coefficient of friction between the object and the fingers (*4*–*6*), which changes over time (*7*). Predicting the motion of rolling contacts similarly depends on these properties, as well as knowledge of local surface geometries (*8*–*10*). Many of these parameters are challenging to measure a priori and even more difficult to measure in real time with previously unseen objects in a changing environment, although some promising sensor advances have been made (*11*).

In this work, we demonstrate highly dexterous manipulation that uses all of the hand's internal surfaces, by a hand that was specifically designed for this purpose. This hand enables manipulation where contacts can seamlessly shift between rolling, fixed, and sliding modes, while object ejection is prevented through caging, and its control does not depend on knowledge of contact forces, friction coefficients, or local geometries. The work was motivated by observing human hand manipulation, which is highly dexterous and uses all surfaces of the fingers and palm, without enforcing fixed contact. Several works in the literature have acknowledged that similar manipulation in robotic hands is rare but leads to increased dexterity when used (*12*–*14*). With the presented work, we make progress toward that kind of dexterity in two main ways. First, we treat the hand-object system holistically and use the observation that the system can be viewed in terms of total summed actuation effort and overall system energy, with corresponding variations in energy based on object location and configuration. Instead of the traditional method of calculating and controlling individual joints and actuators to result in a desired overall system configuration, we drive grasped objects to desired configurations by varying the potential energy in the system, forcing objects to follow energy gradients to a desired state. This is accomplished without having to precisely model contacts or wrenches within the system and is robust to the uncertainties that typically make producing controlled sliding or rolling extremely difficult. Second, we perform these manipulative actions while ensuring that the fingers "cage" the object (*15*–*17*). This physically prevents the object from being ejected during manipulative movements that are generally risky to perform, such as stick/slip transitions.

Over the decades, potentially hundreds of robotic hands have been designed and optimized for every conceivable reason. Hands have been designed for ease of grasping a variety of objects, with many of them based on underactuated mechanisms that compliantly adapt to irregular object geometries (*18*–*21*). Many robotic manipulators have been designed to perform specific dexterous within-hand manipulation tasks, such as object reorientation (*22*, *23*), object translation (*24*), or both (*25*, *26*). Some researchers have successfully automated the design process based on task requirements, like workspace size or high-level manipulation objectives (*27*, *28*), or in other cases for specific manipulation tasks, showing that great manipulation robustness can be achieved by optimizing the controls of the hand in an automated fashion (*29*). In an attempt to design more dexterous hands, researchers have investigated the trade-off between simple but robust manipulators and more complex human-like manipulators using modular grasping strategies (*30*). However, grippers do not need to be mechanically complex to achieve impressive manipulation. A number of simple underactuated hands have been optimized for within-hand manipulation (*31*–*33*).

At the core of every one of these designs is a forward motion model (FMM), a mathematical relationship between actuator inputs and object motions. The FMM enables a prediction of object

Department of Mechanical Engineering and Materials Science, Yale University, New Haven, CT, USA.
*Corresponding author. Email: walter.bircher@yale.edu

motion, given a controllable feature of the system, such as actuator torque or position. Using this prediction, it is then possible to change design parameters within the system and observe changes in the object motion, until a desired behavior is reached. In some cases, the FMM is purely kinematic (*34*, *35*); in others, detailed modeling of contact modes (e.g., rolling, sliding, fixed, and breaking) is included (*36*); and some even model inertial forces and dynamics (*37*).

In this work, we used a potential energy–based FMM that we previously introduced for use with caging-based manipulation with underactuated hands (*38*–*40*). This FMM, known herein as the energy model, does not require direct consideration of force closure or grasp stability and is formulated in this work as a constraint-based minimization. It was inspired by simple linkage-based kinematic gripper models and the energy-based analysis of grasping ability found in (*41*, *42*). A basic model of caging (*43*) was enforced to prevent object ejection and was combined with the object motions predicted by the FMM to form a manipulation metric. This metric was then used to assess the performance of different symmetric hand designs based on every planar open kinematic serial chain finger, made of up to three revolute and prismatic actuators. Last, the design space search yielded a highly dexterous robotic hand that can manipulate objects of many different sizes without ejection; it was constructed, and its dexterity was demonstrated. This hand, called the Model W and shown in Fig. 1, has a prismatic palm and two serial revolute joints per finger; although kinematically very different from a human hand, its motion unintentionally resembles that of thumb–index finger manipulative movements in humans, hinting at a potential design strategy for future robotic manipulators.

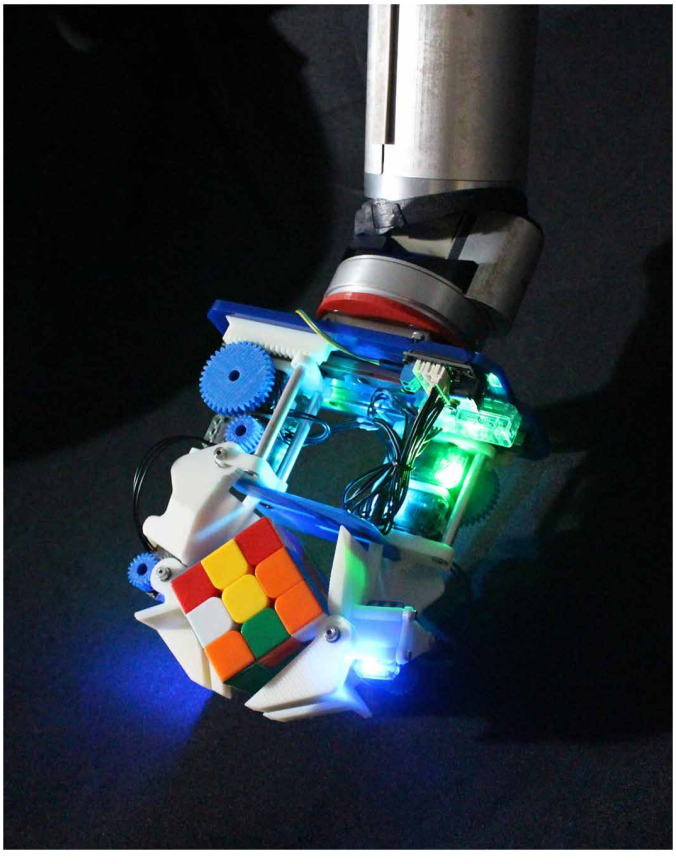


Fig. 1. The Model W dexterous hand was designed to use all inner surfaces of its fingers and palm for manipulation.

RESULTS

Design space search

A total of 6250 unique hand designs based on the 14 kinematic topologies detailed in the "Kinematic topology enumeration" section were simulated by manipulating the 10 distinct objects described in the "Object size, shape, and pose variation" section, and their caging manipulation abilities were graded using the metric described (Eq. 10). The simulation was parallelized in MATLAB and run on 640 cores of the Yale High Performance Computing resource with a total runtime under 12 hours. Selected design spaces are visualized for these topologies in Fig. 2. The results match well with intuition, because simple hands, such as the *R* topology, are incapable of fully manipulating objects in all directions; they can either push or pull objects with respect to the palm depending on their size, but they cannot do both from any point within the workspace. Hands based on the *RRR* topology perform well for small palm widths, because having a palm smaller than the smallest object you intend to manipulate allows you to push it away from the palm, while the fingers are dexterous enough to shift objects in all other directions.

Out of all the unique hands sampled, a few stood out from the rest—capable of caging many objects while achieving high levels of dexterity throughout a large workspace. Specifically, a variant of the *PRR* topology consisting of fixed prismatic joint angle $\phi = 0$ radians and hand proportions of half palm width $p = 0.25$, middle link length $l = 0.43$, and distal link length $d = 0.32$ stood out with the highest manipulability score across all simulated object geometries. To see how these parameters are measured, see Fig. 7. As described in the "Design space variation" section, these values are proportions of total finger length, which includes half of the palm, and can be scaled to any desired physical dimension.

The Model W hand

The physical hand was designed from the optimal design topology and parameters resulting from the brute force search of the design space. The hand was designed to use inexpensive components and a simple design, and its design has been released through Yale OpenHand (an open-source robot hand hardware initiative) in hopes of making it a useful tool to others in the robotic manipulation community. From proximal joint to fingertip, its finger length is 108 mm, and the space between its proximal joints (its palm) can expand from 0 to 72 mm. It is actuated by six Dynamixel XL-320 servos, which are among the least expensive commercially available smart servos—an order of magnitude lower in cost than typical Dynamixel servos—making this hand an affordable open-source option despite having six smart servos. All parts were either commercial off-the-shelf, three-dimensional (3D)–printed from acrylonitrile butadiene styrene (ABS) on a Stratasys uPrint or cut from acrylic using a laser cutter. Each finger is composed of two serial revolute joints connected to a carriage that translates on a linear rail. The distal links of the fingers were made to interdigitate to better approximate the simulated hands. The final hand design is shown in Fig. 3. One of the most challenging aspects of physically designing this hand was allowing the proximal finger joints to come together coaxially (to achieve a zero palm width configuration), with their axis coplanar with the palm. This was achieved using a thin floating palm and cantilevered

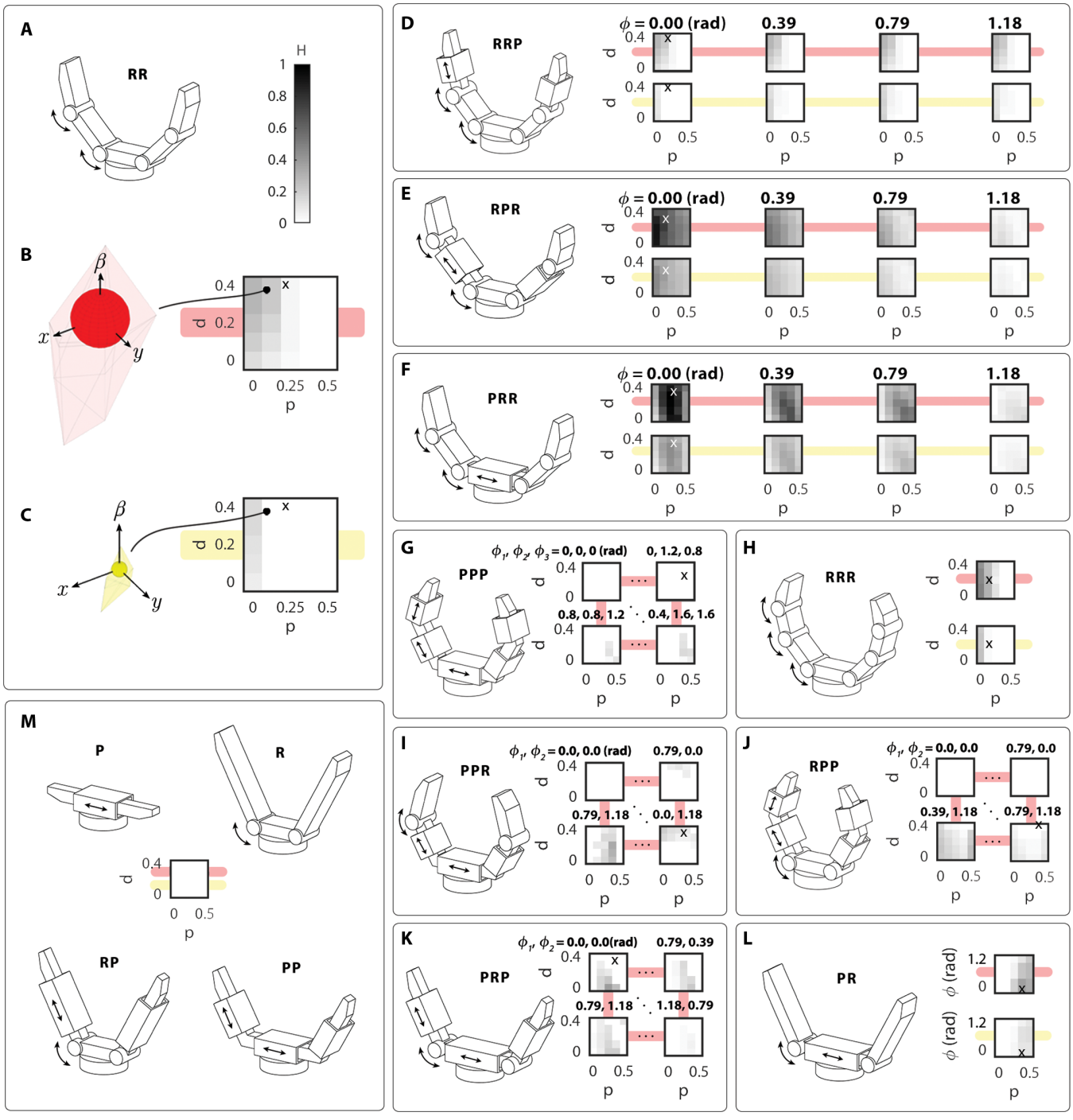


Fig. 2. Simulation results from the design space search. Each shaded square represents a unique hand design. Each hand topology is represented graphically, with arrows drawn to indicate revolute and prismatic joints. Design space plots show how manipulability varies with parameters *p* (palm width), ϕ (fixed prismatic joint angle), and *d* (distal link length) for 14 kinematic topologies. (**A**) The *RR* topology is shown next to the manipulability scale (*H* from Eq. 10), which goes from 0 (cannot manipulate in all directions) to 1 (best manipulability for all objects). (**B**) The manipulability design space for the *RR* topology, calculated from the convex hulls of gradient vectors (left). Red indicates standard manipulability (calculated from $\mathbf{w}_{\mathbf{q}}^{*}$, see Eq. 9) in the rest of the figure. The x on the plot indicates the design parameters corresponding to the pictured computer aided design (CAD) model. (**C**) The caging manipulability design space (calculated from $\mathbf{w}_{\mathrm{cq}}^{*}$). See the "Energy fields" section for more details. Yellow indicates caging manipulability in the rest of the figure. (**D** to **L**) The manipulability for all other viable topologies. Four of 25 plots are shown for topologies with two prismatic joints and 4 of 125 for *PP*. (**M**) The manipulability for topologies that are not capable of fully manipulating objects.

proximal joints, as shown in Fig. 3. Experiments were performed using a variety of control strategies. First, open-loop strategies were implemented, where each joint's servo was simply commanded to a desired set point. In some cases, the servos were immediately commanded to the set point, and in others, they were incrementally stepped to the set point, if a slower speed was desired. In addition, a teleoperated controller was implemented for direct user control of the hand. Last, a visual servoing controller was implemented that used the object's configuration [captured using an object-affixed ArUco marker (*44*) and an overhead camera] to step the actuators of the hand to minimize the error between the object's current configuration and a goal configuration. These controllers are shown in detailed algorithm blocks in the supplementary figures (open-loop control in fig. S1, teleoperation in fig. S2, and visual servoing in fig. S3). It should be noted that none of the controllers had any processes to enforce caging during manipulation—rather, the hand was able to do this naturally by design.

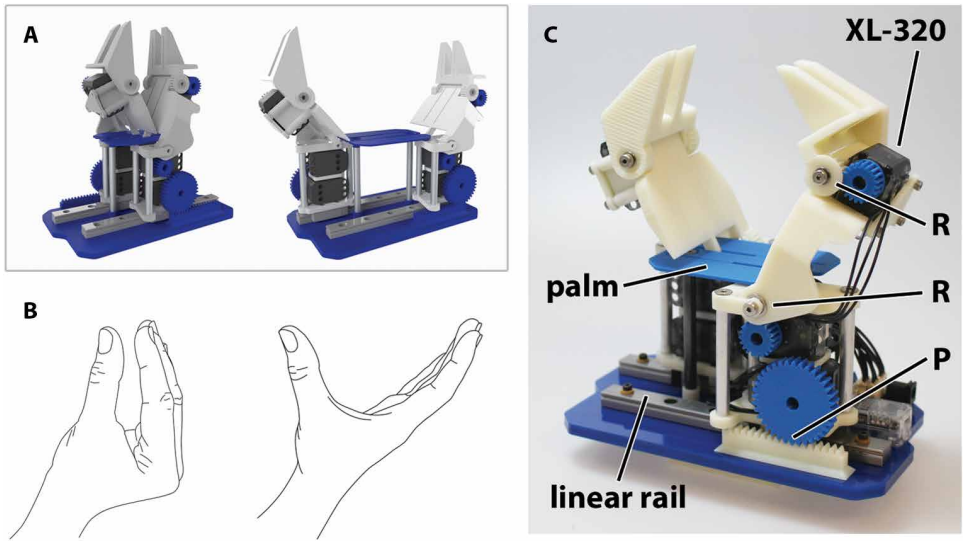


Fig. 3. The physical hand model with labels and a human thumb–index finger manipulative motion for comparison. (**A**) CAD renderings of the Model W in different palm configurations. (**B**) Although kinematically very different, the spreading motion of the Model W's palm is similar to the ability of the human hand to spread the index finger and thumb. (**C**) The physical Model W, with six Dynamixel XL-320 servos. *R* stands for revolute joint, *P* stands for prismatic joint.

Simulation evaluation with test objects

The hand's manipulation capability was first evaluated using 3D-printed test objects. To initially assess the hand, four primitive grasps (power, pinch, left, and right) were manually predetermined for the test objects, and manipulation was achieved by switching between them in an open-loop fashion, meaning that the controller executes the grasp without any feedback about the object's position. The hand was able to perform these motions seamlessly at very high speeds, without ejecting the object, likely due to the inclusion of caging in the metric.

An experiment was performed to assess how well the simulated energy maps predict the actual motion of physical objects. Specifically, we determined how close an object came to the low-energy regions of the workspace, given an actuation input, by recording 360 physical object motion trajectories. This experiment consisted of all combinations of three test objects (T1, T3, and T4 from Fig. 5) manually placed at 30 starting positions in front of the hand (green dots in Fig. 4, B and D) and actuating the hand using the four primitive preprogrammed grasps (power, pinch, left, and right). The hand began actuating when the object was placed within two pixels of the starting location. Then, the motion of the object was recorded using a camera and an ArUco marker (a subset of all object trajectories are shown in yellow in Fig. 4, C and E). Last, the theoretical energy was calculated at the endpoint of each trajectory as a percentage of the maximum energy of the current energy map (for more on how to calculate energy maps, see Eq. 6). The endpoint energy percentage of each trajectory was recorded and is shown in the histogram in Fig. 4A. Most of the trajectories resulted in final object positions in regions with less than 10% of the maximum theoretical energy.

Workspace evaluation with test objects

A simple visual servoing–based controller was implemented that interpolated between the four primitive grasps while simultaneously

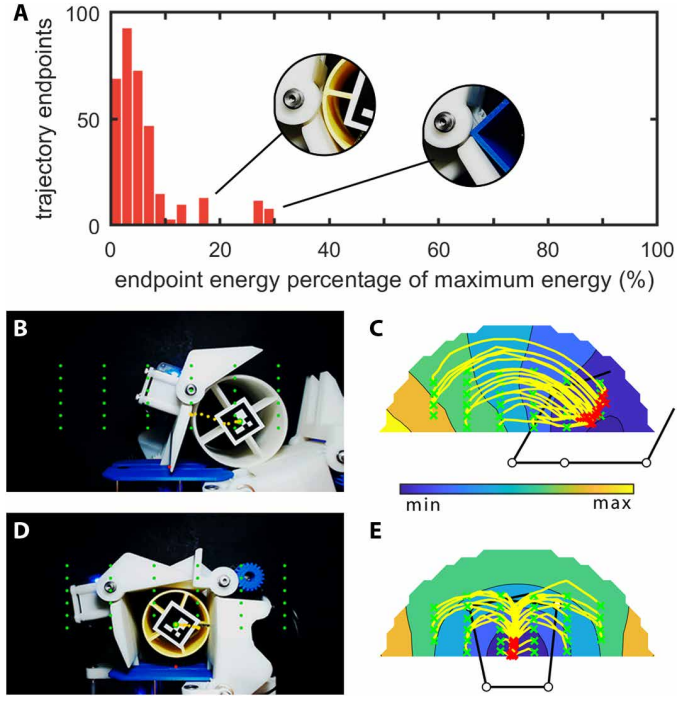


Fig. 4. The Model W experimentally manipulates objects to low-energy regions of the workspace as predicted by simulated energy maps. (**A**) A histogram showing the distribution of object trajectory endpoint energy percentages of all 360 trials. The distribution shows that objects move to low-energy (<10%) regions of the workspace as predicted by simulated energy maps and that energy maps are actionable. The small peaks near 20 and 30% energy correspond to "pinch grasp" trials in which the object became stuck in the unsimulated physical geometry of the distal joints. (**B** and **C**) Left: Photograph of the "right" grasp trials with object T3 with grid of starting locations (green markers). Right: Corresponding energy map, object trajectories (yellow lines), and object ending locations (red markers). (**D** and **E**) The same images for the "power" grasp trials with object T1.

limiting the torque output of each motor (see fig. S3 for details). Limiting the torque allowed the motors to be driven to stall without worry of overheating, enabling manipulation of objects of all sizes with the same set of predetermined grasp primitives.

Using this controller, objects were manipulated on a support plane to randomly generated goal waypoints in front of the hand, and the system was allowed to run continuously until 200 waypoints were reached per object. The resulting workspaces of four test objects are shown in Fig. 5. The controllable translational workspaces are shown for all objects, and an additional plot showing the controllable rotational workspace is provided for a square object. The workspaces are large compared with the hand, most notably along the axis normal to the palm. This is partly because the hand is able to perform a power-to-pinch transition, a challenging task for many prehensile manipulators.

The visual servoing controller was then modified to control object orientation rather than position, and an experiment was performed that attempted to manipulate the square test object T4 to 180 randomly generated goal orientations between 0° and 90°. The results of this experiment are shown in Fig. 5E, with reached orientations shown in red and unreached orientations shown in black. The controller was allowed 100 cycles of the loop (100 motor commands) to reach each goal orientation, and unreached goals occurred when the controller could not reach the goal in this time.

Real-world manipulation scenario

The Model W was fitted to a 7 degrees of freedom (DOF) WAM Arm (Barrett Technology), and a variety of manipulation tasks were performed using the Yale-CMU-Berkeley (YCB) Object Set (*45*) to demonstrate its performance in real-world tasks. First, the hand was used to grasp different objects, showing the range of object sizes that can be effectively picked by the hand. The hand can easily grasp very small items, like dice, as well as very large objects like the widest dimension of a cracker box (210 mm) or a mini soccer ball (140-mm diameter). The range of graspable objects is largely due to the prismatic nature of the palm, allowing the fingertips to come completely together or to spread apart more than 300 mm.

Next, a series of open-loop manipulation tasks (see fig. S1 for details) were performed on a tabletop to demonstrate within-hand manipulation of real-world objects. As a first manipulation task, both rigid and soft cubes were continuously reoriented on a support surface. The first one, the puzzle cube from the YCB Object Set, was easily reoriented using a sequence of preprogrammed grasps. The second object, a hand-made deformable knitted cube (50 mm) containing

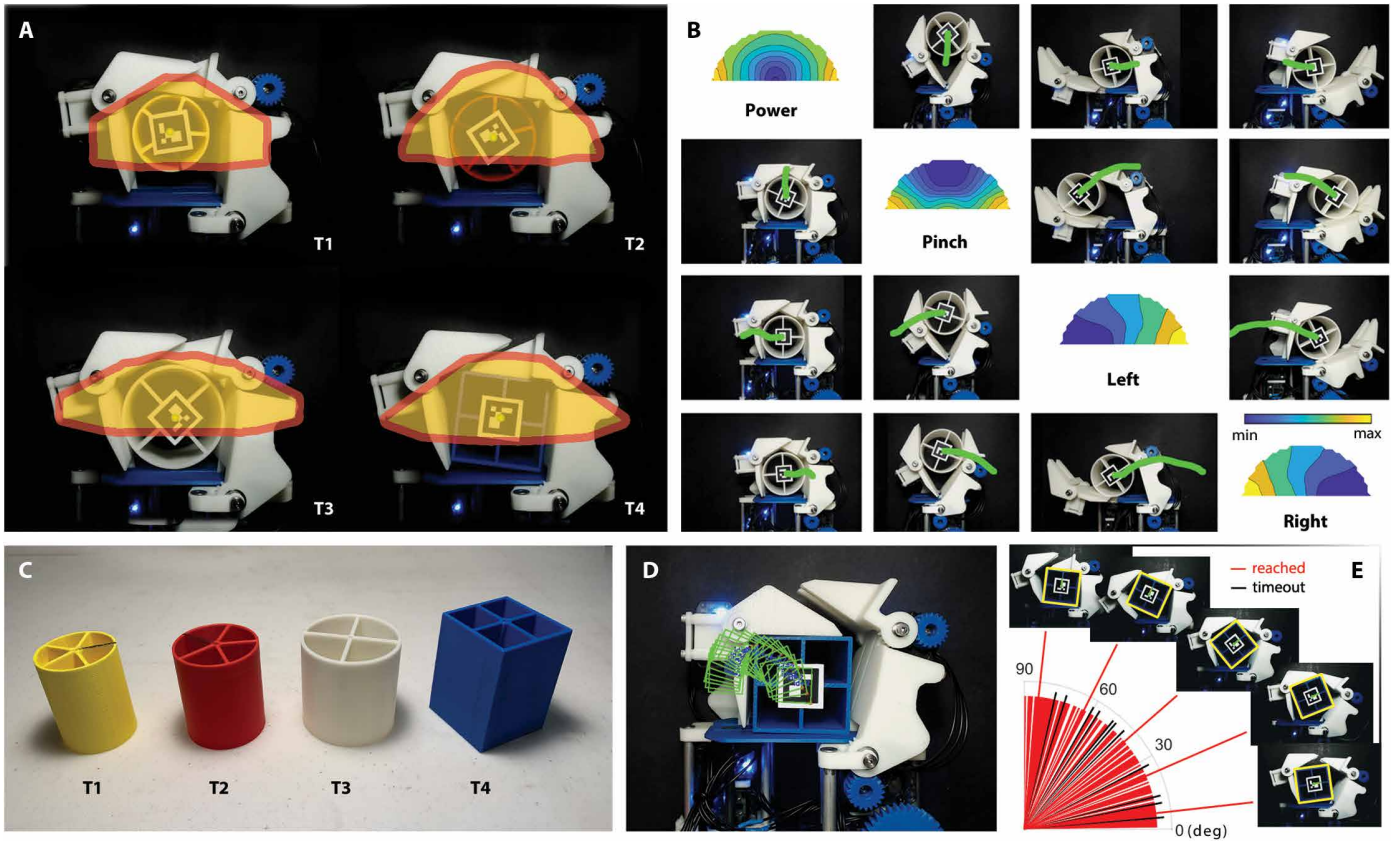


Fig. 5. The Model W has a large fully connected workspace and can continuously rotate asymmetric objects. (**A**) The representative workspaces with the four test objects shown in (C). (**B**) The hand's ability to translate objects in every direction. Each photo shows the hand manipulating the T3 object from the diagonally listed direction in its row and to the direction in its column. For example, the photo in the first row and second column shows the hand manipulating from a power grasp to a pinch grasp. (**C**) The test objects used during benchtop experiments. (**D**) An example of the hand continuously rotating the square test object. (**E**) Data from the random goal orientation servoing experiment showing goal orientations reached (red) and orientations where the controller timed out before reaching the goal (black). Select example photographs are shown from the experiment, and the object is outlined in yellow for clarity. The objects are supported by a tabletop in all tasks.

beans, was also successfully reoriented in a continuous fashion, albeit much slower because of less efficient force transmission due to deformation. The next manipulation scenario performed was to reorient, grasp, and squeeze a mustard bottle. The bottle, which was placed on a surface in front of the arm, was rotated about 90° within hand to facilitate a proper squeeze, grasped, and lifted above a plate; last, mustard was squeezed onto a plate. After this task, the plastic YCB orange was repeatedly shifted from power to pinch grasps and vice versa. Photographs taken during the completion of these tasks are shown in Fig. 6.

A series of teleoperated tasks (see fig. S2 for details) were also performed using a shape sorting cube toy, motivating hand dexterity over arm dexterity, and demonstrating within-hand manipulation and controllability, despite the rather high number of individual actuators in the hand. During these tests, the hand successfully and repeatedly manipulated a set of blocks into their respective holes.

Last, the hand was tasked with manipulating multiple freeform objects at once to assess its ability to handle unstructured objects on a support plane. First, we successfully manipulated both a squash ball and a golf ball simultaneously, juggling them around a common center of rotation within the hand. Not only are the balls of different sizes, but their surface textures are very different from each other—the golf ball being hard and dimpled, and the squash ball being rubbery and compliant. Even with these differences, the hand was easily able to perform this task in both directions (clockwise and counterclockwise juggling around a common center of rotation). As a final assessment of the hand's abilities, we manipulated Chinese Baoding balls, which are sometimes used to improve human dexterity after surgery. Using the same controller from the squash/golf ball manipulation, the hand was once again able to juggle these balls, even though they are much larger in diameter. We then programmed the robotic arm to continually travel to different waypoints in the plane while performing this manipulation, applying constant external disturbances to the balls. Despite the added disturbances to the system, the manipulation continued successfully, and the balls were not lost from the grasp. By creating a cage during manipulation, objects were much less susceptible to being ejected from the hand, and their stability in the grasp was no longer dependent on parameters related to contact position, forces, or mode. Still photographs from these experiments are shown in Fig. 6. All of these demos and more are shown in the accompanying main text video (Movie 1).

DISCUSSION

Assumptions, limitations, and approximations

This work would not have been possible without many simplifying assumptions and approximations. As such, we feel that it is very important to discuss them candidly, because their limitations and implications will undoubtedly help to shape future work in this area.

First, the scope of this work was intentionally limited to planar symmetric manipulators with up to three actuated elements per finger. The reason for this was primarily computational. Adding more

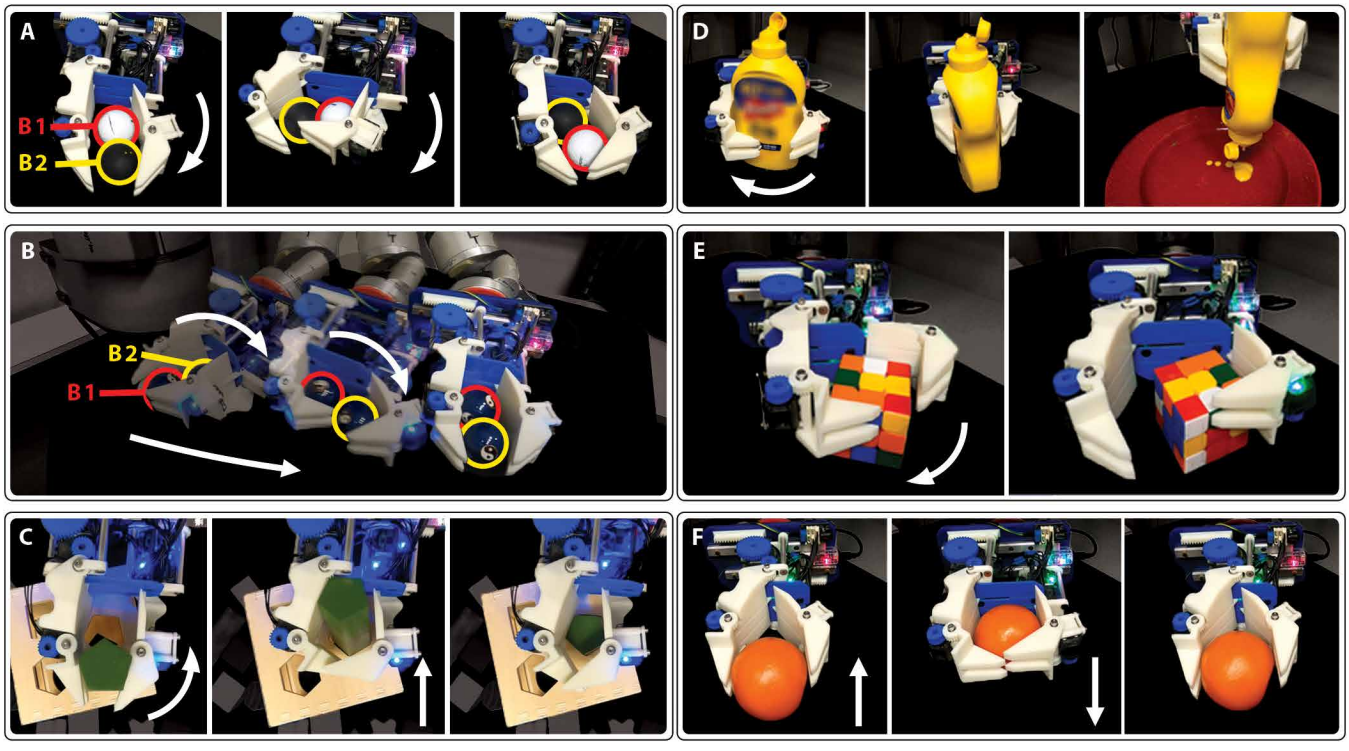
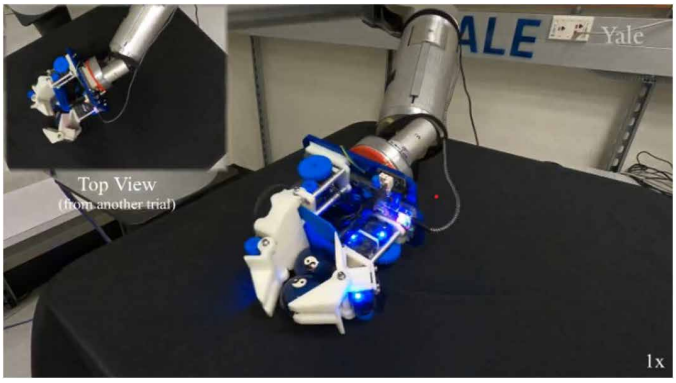


Fig. 6. The Model W can manipulate multiple objects at once, can be controlled using teleoperation, and can reorient and perform pinch-to-power transitions with real-world objects. (**A**) A golf ball (B1) and a squash ball (B2) are rotated about a common point of rotation. (**B**) Two Chinese Baoding balls (B1 and B2) are rotated about a common point of rotation within the hand while it is simultaneously moved to different waypoints within the plane, demonstrating how caging can prevent object ejection during manipulation with external disturbances. (**C**) The hand manipulates different wooden blocks into holes of matching shapes using teleoperated control. (**D**) A bottle of mustard is reoriented on a table, grasped, and squeezed. (**E**) A puzzle cube is rotated continuously within the hand. (**F**) An orange is transitioned from pinch grasp to power grasp and back again. The objects are supported by a tabletop in all tasks. The background of all photographs is darkened to better highlight the hand-object system.



Movie 1. Summary video of the Model W. An overview of the dexterous capabilities of the Model W.

actuators to each finger not only increases the number of possible kinematic topologies but also increases the number of unique combinations of actuation inputs to the system exponentially and thus the number of required energy maps that must be calculated. For this reason, it is likely impractical to use this technique for systems having greater than a few actuators in series. Extending this work to asymmetric hands is not as computationally taxing as increasing the number of actuators per finger, because it simply requires pairwise manipulability scoring of all unique finger topologies already presented. No additional energy maps would need to be calculated; rather, all unique pairs would simply have their energy maps summed and could be scored accordingly.

Extending this work to 3D raises another challenge. Specifically, the constraint in Eq. 6 was formulated for planar problems and cannot directly be extended to a third dimension. When imagining that spatial fingers could include actuated segments that move out of plane with one another, the kinematic constraints will become less straightforward and will likely slow down the computation speed of the system energy. Future work into more advanced kinematic topologies (featuring spherical joints or even out-of-plane prismatic joints) depends on the reformulation of this constraint. It is likely that basic collision detection–type constraints will suffice in the spatial case, but they will require more computation time.

Even in the plane, this constraint has a number of limitations that should be considered. First, it only works with fully convex finger geometries, meaning none of the revolute joint or prismatic joint angles hyperextend and are negative. This was not a problem for us, because we assumed that all fixed prismatic joint angles and all actuation inputs were positive, but it does prevent more exotic actuation schemes and prismatic joint designs. Ideally, a more general kinematic constraint could be formulated to avoid this issue, which may also enable simulation of advanced spatial topologies.

Aside from this constraint, there are other large assumptions governing the energy maps that are worth the discussion. First of all, the model does not take friction into account in any way. This is obviously quite an unrealistic assumption in the real world, but we still see that the hand successfully drives objects to low-energy regions despite this. We do not know how much friction can be introduced between the hand and the object before the object can no longer be manipulated to low-energy regions, but we plan to investigate this in future work. Furthermore, there is no guarantee that the theoretical wrenches can physically be applied to an object for a known actuation input. The model also assumes that motors can be backdriven or displaced, resulting in an increase in potential energy. Applying this assumption with physical motors means choosing actuators without heavy gearing or using position control with intentionally low torque limits. Speed of operation may play a role in how well the model works in practice, because dynamic effects may dominate the motion with heavier objects. To truly understand the advantages of this model, we must work to characterize these limitations in a quantitative fashion.

Design space search

The results in Fig. 2 highlight a number of key takeaways that can be used as general guidelines for the design of future manipulators. First, hands with proximal revolute joints perform best with small palm widths, as illustrated by topologies *RR*, *RRR*, *RPR*, and *RRP*. This is because these topologies are unable to perform power-to-pinch transitions with objects smaller than their palm width. Conversely, topologies with prismatic palms avoid this problem, because they are all capable of a wide range of palm widths. This explains the stark visual difference between the best performing topology *PRR*, which has a fully shaded design space, and others beginning with a revolute joint that only has a shaded vertical bar for small palm width designs. Next, fully symmetric prismatic topologies are poorly performing in part due to their inability to easily reach each finger across the palm, limiting the shared workspace of both fingers, and also because they can only apply motion in a fixed direction. Intuitively, it would seem that *PPP* should be able to manipulate very well. However, high-quality manipulation is only possible

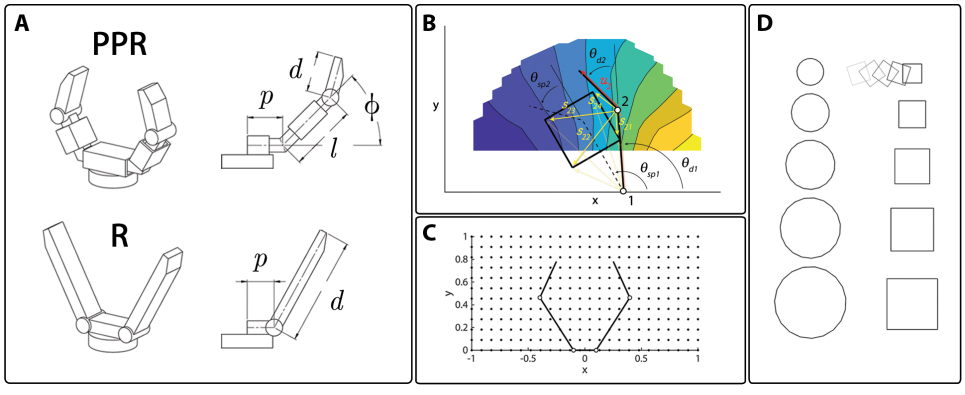


Fig. 7. Simulation design parameters visualized. (**A**) Two example topologies, *PPR* and *R*, are shown here with labeled dimensions. The distal-most link length is *d*, the angle of prismatic joints is ϕ, middle links are of length *l*, and the palm width is *p* (one-half of the palm is considered to be part of each finger). (**B**) A diagram illustrating key vectors used in the constraint in the optimization program described in Eq. 6 and also showing how joint angles are measured. (**C**) The objects were simulated at all positions shown in the 24-by-12 grid (example *RR* hand is shown for scale). (**D**) The 10 simulated objects are shown here at the same scale as the hand and grid in (C). Five distinct sizes of square objects and circular objects were simulated. The square objects were simulated at six orientations each to capture the hand's ability to reorient them.

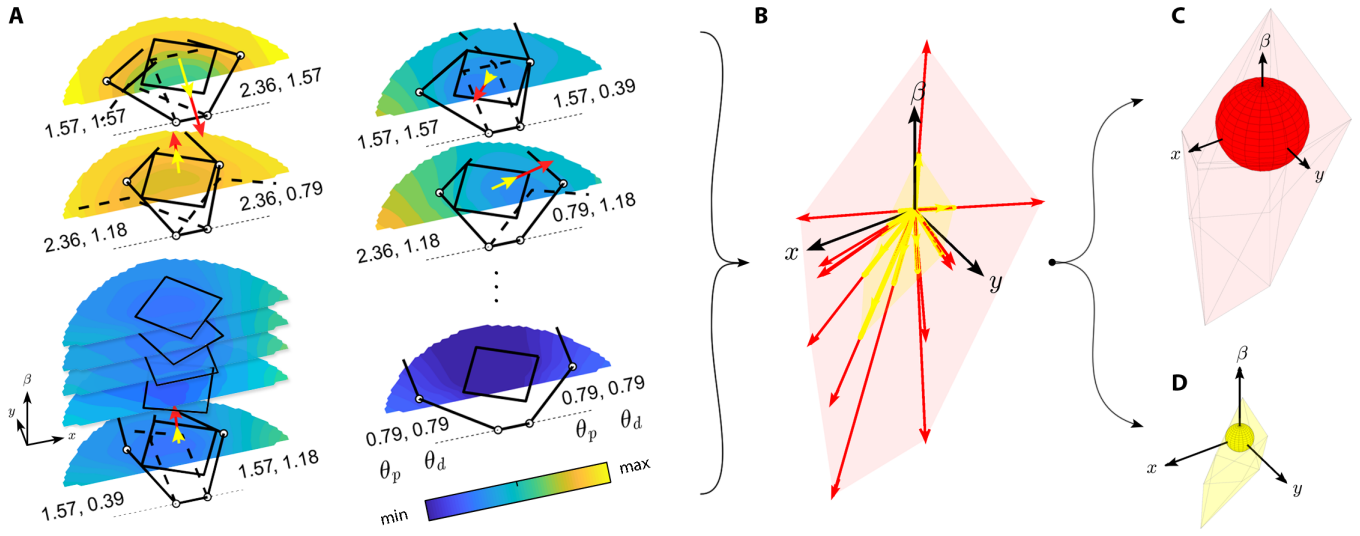


Fig. 8. Manipulability is derived from the convex hull formed by gradient vectors of the system's potential energy scalar fields. (**A**) Each contour plot shows the potential energy and gradients for a square object and a symmetric *RR* hand, given a distinct set of actuation inputs. The bottom left stack of contour plots (next to the coordinate frame) illustrates that the energy space is 3D for a planar object (other than a circle), and the wrench vectors shown are also in three dimensions. For the sake of clarity, the other contour plots show only a single slice of the energy space, corresponding to the illustrated orientation of the square. The dashed lines represent the commanded finger positions θ_{sp}, given the actuation set points for each joint (in radians) listed below each contour plot, and the solid lines represent the realized finger positions based on the displaced joint angles θ_d, were they to close around the object in its fixed position. The red vectors emanating from the center of the object represent the gradients of the potential energy fields at the object's center point. The yellow vectors are the red vector scaled by the strength of the basic cage created around the object by the fingertips, a value from 0 (no cage) to 1 (fingers interdigitate). (**B**) The manipulation metrics are calculated by joining the tails of all vectors, calculating the convex hull of their tips, and finding the radius of the largest origin-centered ball contained within their hull. Physically, this radius is proportional to the minimum wrench the fingers can potentially apply to the object in any direction (in the object's configuration space *x*-*y*-β). (**C**) The manipulability (without caging) is represented by the radius $\mathbf{w}_{\mathbf{q}}^{*}$ of the red ball centered at the origin and (**D**) the caging manipulability by the radius $\mathbf{w}_{cq}^{*}$ of the yellow ball.

with asymmetric designs of this topology, because you need one finger to push the object away from the palm, while the other is able to draw it closer—one finger is not capable of both because frictional forces are not considered in this model.

Another takeaway is that modifying the metric to include caging moderately lowers the scores of many hands. In other words, the caging design spaces (yellow) shown in Fig. 2 are perhaps less intuitive than the corresponding design spaces that do not include caging (red). Many topologies are actually better at manipulation than conveyed by the caging manipulation plots but simply cannot manipulate while simultaneously caging the object. This is seen by comparing the corresponding red and yellow design space boxes in any topology that shows both, where the plots corresponding to "no caging" always have more pronounced shading than "caging." Intuitively, this makes sense, because the caging metric will leave the manipulation wrenches the same at best or scale them back at worst. With this in mind, in Fig. 2, the caging manipulation design spaces can be imagined as slightly lighter versions of the manipulation design spaces for any topologies where caging design spaces are not shown (any topologies with more than one prismatic joint). When considering hand designs based on manipulation ability alone, both *PRR* and *RPR* were the top performers, and the Model W was designed on the basis of the parameters of the best *PRR* hand.

Experimentation

One limitation of this work is the grasping strength of the hand. Because the hand was designed with inexpensive hobby motors and because the finger surfaces were left as unfinished low-friction ABS plastic to facilitate easy rolling and sliding, the grasp strength of the hand was quite low. A second iteration of this hand could benefit from both stronger motors and from actuated friction pads that can retract into the fingers on command such as in (*46*).

Despite the rather low grasping force, this hand was highly capable of manipulating a wide range of objects. Specifically, the hand was easily able to perform continuous object rotation, pinch-to-power transitions, simultaneous manipulation of multiple objects at once, manipulation in the presence of external disturbances, manipulation of soft objects, and object position and orientation servoing within the hand. The hand demonstrated very robust manipulation under the presence of continually changing and unpredictable contact conditions (e.g., multiple sliding and rolling contacts between the object and the hand that were constantly broken and reestablished).

Often, the same open-loop strategies worked for multiple objects of different sizes, due to the low torque threshold set for the motors. In a sense, the torque limits produced a compliant finger behavior: A large enough force applied to the object by one finger could backdrive the opposing finger, depending on its relative mechanical advantage. Speaking purely qualitatively about our experience using the hand, we can say that it almost never lost an object from its grasp, despite many situations where the object was briefly not caged by the hand. We thought that this was notable, because we did not explicitly control the hand to do this, and rather, it suggests that our design metric worked as intended. The inclusion of caging in the metric likely increased the range of experiments that we were able to attempt without controller modifications, because we rarely had to deal with object ejection, and could therefore be more adventurous

Table 1. List of notations used in the paper.

U	Potential energy
τ	Torque generated by motor
θ_d	Displaced joint position
θ_{sp}	Motor set point
n	Number of actuators
K	Transmission ratio
R	Rotation matrix
u	Joint position vector in Cartesian space
P	Set of Cartesian positions of all joints in a finger
${}^aT_b$	Homogenous transformation matrix from frame a to b
N	Dimension of system, 2 for planar, and 3 for spatial
S	Set of all object boundary points
s	Object boundary point in Cartesian space
U^*	Minimum energy
q	Object pose
U	Minimum energy-valued scalar field
Q	Number of distinct object poses
γ	Vector field
w	Wrench vector
x, y, β	Planar object pose coordinates
Θ	Number of distinct actuation inputs
W	Set of vectors at single object pose for all actuator inputs
H_O	Hand caging manipulability score with object O
R, P	Abbreviations for revolute and prismatic joints
p	Palm width
d	Distal link length
ϕ	Prismatic orientation with respect to previous link
l	Link length

in the manipulation strategies attempted through control, regularly finding success in physical demos during the first few tries.

MATERIALS AND METHODS

In this section, we describe the formulation of our energy-based FMM. In short, we compute a hand-object workspace, such that each point in the workspace has an associated energy value. We term this workspace an energy map. There is one energy map (in the form of a 3D scalar field where the dimensions represent the object's planar position and orientation) associated with each commanded actuation input to the system. We compute the gradient of each energy map, resulting in a vector field for each actuation input. Each vector in the field corresponds to the direction and magnitude of motion that can be realized by the object at that configuration in the workspace, given the corresponding actuation input. Thus, for a given object position and orientation in the workspace of the hand, there will be many distinct directions of feasible object motion, corresponding to the collection of all gradient vectors from all possible actuation inputs to the system.

Forward motion model

This theory is based on the idea that with enough mechanical work, a position-controlled motor can be displaced from its commanded set point. Equivalent in magnitude to the work put into the motor, the potential energy gained by a motor during such a displacement is

$$U = \tau_m(\theta_d - \theta_{sp}) \quad (1)$$

where θ_{sp} is the commanded set point, τ_m is the torque generated by the motor (assumed to be a source of constant torque), and θ_d is the displaced position of its shaft (see Table 1 for descriptions of all symbols). A displacement can be caused by an external disturbance that cannot be resisted by the torque generated by the motor. This displacement may occur on the output of a transmission, such that it is propagated backward to the motor. In the case of a transmission, we can relate the observed output displacement to the motor shaft displacement by $\theta_{out} = K(\theta_d - \theta_{sp})$, where K is the transmission ratio, $\theta_d - \theta_{sp}$ is the relative displacement of the output, and θ_{out} is the displacement of the motor shaft. If multiple actuators are displaced, their respective energies are simply summed to compute the system's total energy. The total increase in a system's potential energy due to a displacement of multiple actuators with unique transmissions is

$$U = \sum_{i=1}^{n} \tau_i K_i(\theta_{d_i} - \theta_{sp_i}) \quad (2)$$

where n is the number of actuators in the system. As an example, imagine a simple planar finger modeled as a two-link revolute serial chain, with a motor at each joint. If each joint of the finger is commanded to a known set point and an object is forcibly brought past the point of contact with both links of the finger, both motors can be backdriven simultaneously, and both joints will be displaced, due to the work done on the finger by the forcible placement of the object.

System kinematics

This displacement in motor position can easily be mapped to displacements in Cartesian positions of joints of the finger using forward kinematic transformation matrices, constructed from Denavit-Hartenberg parameters. The Cartesian position of a frame affixed to joint k of a n link serial chain linkage with respect to a fixed global frame is $\mathbf{u}_k$ extracted from the transformation matrix ${}^0T_k$ where

$${}^0T_k = \begin{bmatrix} R_k & {}^0\mathbf{u}_k \\ 0 & 1 \end{bmatrix} = \prod_{i=1}^{k} {}^{i-1}T_i(\theta_{sp_i}), k \le n, {}^0\mathbf{u}_k \in \mathbb{R}^N \quad (3)$$

such that R_k is the rotation matrix associated with the kth joint, ${}^0\mathbf{u}_k$ is the joint position vector for the kth joint with respect to a fixed global frame, and ${}^{i-1}T_i(\theta_{spi})$ is the homogeneous transformation matrix that transforms points from the link $i-1$ affixed frame to

link i. We assume that joint-affixed frames follow the standard Denavit-Hartenberg convention, with the z axis aligned with the joint axis, and we also set positive joint displacements such that they result in a finger becoming more closed. We also assume that revolute joint angles are measured with respect to the previous link and cannot hyperextend past the previous link to a negative angle (in other words, $\theta_i \geq 0$). For each finger, let $P_{\theta_{sp}}$ be the set of Cartesian positions of all joints with respect to the previous joint (${}^0\mathbf{u}_k - {}^0\mathbf{u}_{k-1}$ from Eq. 3) at the commanded actuation set point θ_{sp} such that

$$P_{\theta_{sp}} = \{{}^{k-1}u_{\theta_{sp},k} \,|\, k = 1,\ldots,n\}, {}^{k-1}\mathbf{u}_k \in \mathbb{R}^2 \quad (4)$$

An object in pose q can be represented by a set S of m boundary points with respect to a joint-affixed frame i such that

$${}^iS_q = \{{}^is_{q,ij}s_{q,ij} \,|\, i = 1\ldots n, j = 1,\ldots,m\}, {}^i\mathbf{s}_j \in \mathbb{R}^2 \quad (5)$$

where $\mathbf{s}_j \in \mathbb{R}^2$ for planar objects.

We imagine an object fixed in space displacing the link of a finger by a known amount in Cartesian space, but we actually wish to know the displacement of the motor driving that link of the finger (the displaced joint angle θ_d) to compute the gained potential energy. This implies the use of inverse kinematics—to find a joint angle that produces a desired link position, such that contact occurs with the object somewhere along the finger. However, because we are not limiting this work to fingertip manipulation, we do not know precisely where contact occurs between the link and the object, and instead of directly calculating candidate displaced joint angles using inverse kinematics, we frame the problem of determining displaced joint angles as a mathematical program that minimizes the system's total energy, using a nonlinear constraint based on the homogeneous transformation matrices in Eq. 3. Specifically, we formulate a constraint that prevents any part of a finger both from penetrating an object and from passing from the finger's open configuration through the object to the other side.

Energy minimization

Because we are interested in finding the displaced configuration of the hand that corresponds to its lowest energy state, we formulate an optimization problem to minimize the energy. To compute the potential energy gained by a position-controlled serial chain finger that is displaced by an object of known geometry and configuration, we compute the minimal potential energy of the system according to the following constraints

$$\begin{aligned} U^*_{\theta_{sp},q} &= \min_{\theta_d} \sum_{i=1}^{n} \tau_i K_i(\theta_{d_i} - \theta_{sp_i}) \\ s.t.\ 0 &\leq \theta_{d_i} \leq \theta_{sp_i}, i = 1\ldots n \\ u_{i_x}s_{ij_y} - u_{i_y}s_{ij_x} &\geq 0, j = 1\ldots m \end{aligned} \quad (6)$$

where $U^*_{\theta_{sp},q}$ is the minimum system energy for object pose q and commanded actuation set point θ_{sp}. The displaced finger configuration corresponding to the minimum energy of the system is represented by the vector of displaced joint positions θ^*_d; u_{i_x}, u_{i_y}, s_{ij_x}, and s_{ij_y} are the x and y components of the ith $\mathbf{u}$ and ijth $\mathbf{s}$ vectors (joint positions and object points); P_{θ_d} is the set of all Cartesian joint positions calculated from displaced joint angles θ_d; and θ_{sp_i} is the commanded set point of the ith joint. The constraints $\theta_{d_i} \geq 0$ and $\theta_{d_i} \leq \theta_{sp_i}$ represent that the displaced joint positions must be greater than or equal to zero, usually due to mechanical hard stops, and must be less than or equal to the commanded set points θ_{sp} because the object cannot pull the finger past where it is commanded to go. The final constraint is $u_{i_x}s_{ij_y} - u_{i_y}s_{ij_x} \geq 0$ for right-hand fingers and $u_{i_x}s_{ij_y} - u_{i_y}s_{ij_x} \leq 0$ for left-hand fingers and accounts for which side of the object the finger is on. The constraint is derived from the sign of the cross product of two vectors originating at a joint, with one directed to the subsequent joint (or fingertip) and the other to a point on the object (as illustrated in Fig. 7B). When the sign of this constraint drops below zero, it is physically equivalent to the finger passing through this point on the object or having passed completely through the object to the other side. Visually, this occurs when the tip of one vector passes across the other (remembering that their tails are connected). Although more general contact constraints exist for this problem, this one was selected because it is fast to compute and represents the problem well in this planar situation. Using the logic from Eq. 2, we can simply solve the mathematical program in Eq. 6 once for each finger and sum their minimum system energies to obtain the total hand-object potential energy at each configuration of the system.

Energy fields

The solution to the energy minimization yields the displaced, minimum energy state of the system corresponding to one single pose of the object, given commanded finger actuation inputs θ and fixed object pose q. This can be extended to the set of all possible poses of the object for that actuation input

$$\mathbf{U}_\alpha = \{U^*_{\alpha,q} \,|\, q = 1,\ldots,Q\} \quad (7)$$

resulting in a scalar field $\mathbf{U}_\alpha$ that, when visualized, lends good intuition about how the hand will move the object. In addition to considering how the hand is displaced by the forcible placement of the object in a specific pose, we can also consider how the object will be displaced by the actuation of the hand, should we drop our assumption that it is fixed firmly in place. As an example, a hand's translation capability with a square object is illustrated by the contour plots showing the scalar field $\mathbf{U}_\alpha$ in Fig. 8, where the contour color is the potential energy magnitude at that specific position in the hand-object workspace. Whereas we were initially only concerned with the lowest energy configuration of a hand, we now imagine that the object is released from its fixed configuration and free to move along with the hand and that the entire system will settle to its lowest energy configuration. From the contour map, it is clear to see where the object will tend to move—toward the lowest-energy region in the hand's reachable workspace.

Gradient of energy map

The gradient of the potential energy scalar field with respect to the object's planar pose coordinates x, y, and β results in a distinct vector field

$$\gamma_\theta = -\nabla_{x,y,\beta}\mathbf{U}_\alpha \quad (8)$$

for each scalar field $\mathbf{U}_\alpha$ corresponding to a distinct actuation input α. Each vector field γ_θ consists of pose q–dependent planar wrench vectors $w_q \in \mathbb{R}^3$ of the form $w_q = [w_x, w_y, w_\beta]$. Each of these vectors is the net wrench that can potentially be imparted on the object or, more generally, the wrench that could potentially be realized by the object at pose q given the actuation inputs α from the hand. We qualify these

statements as wrenches that could "potentially" be realized, because these wrenches are calculated in the absence of friction, assuming ideal hand and object geometry, and ideal actuation. As such, the actuation that produces them is a necessary, rather than sufficient, condition for the physical existence of these wrenches. The set of all vector fields $\Gamma = \{\gamma_1, \ldots, \gamma_\Theta\}$ calculated over the set Θ containing all possible actuation inputs can be used to evaluate a specific hand's manipulation capability throughout the workspace of a hand-object system.

Specifically, if we consider a single object pose q with respect to the hand, we can collect the set of vectors $\mathbf{W}_q = \{w_{q,1}, \ldots, w_{q,\Theta}\}$ (one vector from each field γ_θ), where each vector corresponds to a possible actuation input to the hand while the object is at pose q. The span of this set of vectors in $x - y - \beta$ space physically represents the range of wrenches that could potentially be applied to the object. More generally, it tells us how well the hand can manipulate the object from its pose in the workspace, and it is at the core of the manipulation metric used in this work.

Manipulation metric

Given the set of net wrench vectors $\mathbf{W}_q$ that can be imparted on an object at a known pose, we can determine the possible directions of motion that can potentially be realized by the object. This is illustrated visually by the sets of red and yellow vectors in Fig. 8. The convex hull of $\{w_{q,1}, \ldots, w_{q,\Theta}\}$ is Conv($\mathbf{W}_q$) and is the set of all allowable configurations of net wrench application to the object. We want to find the maximum radius ball that can be inscribed in Conv($\mathbf{W}_q$), because the radius of this ball is equivalent to the maximum net wrench magnitude that could potentially be applied to the object in any direction.

In other words, the radius of this ball allows us to compare how well the hand can manipulate the object in one pose relative to other poses. The greater the radius, the larger the wrench we can exert on the object in any direction. To compute the maximum radius ball, we use the hyperplane representation of a polytope formed by Conv($\mathbf{W}_q$). Recall that any polytope is defined by the intersection of hyperplanes $\langle h_i, w\rangle \leq b_i$. See (*47*) for more of this theory and (*48*) for the theory behind producing any polytope representation given vertices, as we have in Conv($\mathbf{W}_q$). Last, given a hyperplane representation, note that the radius $w^*_{\mathbf{q}}$ of the maximum radius ball may be computed with the quadratic program

$$w^*_{\mathbf{q}} = \max_{w} \|w^2\| \tag{9}$$

$$s.t. \langle h_1, w\rangle \leq b_1$$

$$\langle h_2, w\rangle \leq b_2$$

$$\vdots$$

$$\langle h_n, w\rangle \leq b_n$$

A visual example of the largest origin-centered ball bounded by the convex hull of the gradient vectors is shown by the red and yellow balls in Fig. 8.

Caging metric

Although a larger radius tells us that we have greater control authority over the object at a given pose, it tells us nothing about the likelihood of object ejection while continuing manipulation in a given direction. Ideally, we would like to design a hand that can manipulate an object in any direction from any position in the workspace while safeguarding against object ejection. To quantitatively consider this trade-off between manipulability and ejection prevention, we scale each of the vectors in $\mathbf{W}_q$ by a distinct value between 0 and 1, derived from a basic representation of how caged the object is, given the hand's configuration under actuation input θ_{sp}. If the hand does not cage the object at all (meaning that the smallest distance between the fingers through which the object could escape is larger than the object itself), the vector is multiplied by 0, meaning it is not safe to proceed actuating in that direction. If the fingers completely cage the object (meaning there is no gap between the fingers), the vector is multiplied by 1. When a gap exists between the fingers, but it is smaller than the diameter of the object, the vector is scaled by a value between 0 and 1 based on the size of the gap normalized by the size of the object. Now, using this set of scaled vectors, the same procedure is used as in Eq. 9 to find the largest inscribed ball inside of the caging convex hull, resulting in the maximum caging manipulation radius $\mathbf{w}^*_{\text{cq}}$ and illustrated by the yellow vectors and ball in Fig. 8. The caging manipulability score H_O for a hand-object system is then

$$H_O = \frac{1}{U_{\max} Q} \sum_{q=1}^{Q} \mathbf{w}^*_{\text{cq}} \tag{10}$$

where O is the object being manipulated, Q is the total number of reachable object poses where both fingers make contact with the object, and $U_{\max}$ is the maximum potential energy that the system could theoretically see given the actuation limits of the hand. This acts as a scaling factor that accounts for the fact that gradient vectors grow longer (and thus the radius of the largest ball increases) as you add more actuators to the system (you are able to exert a larger wrench on the object given more actuation energy). In simple terms, H is just the average caging manipulability that can be had per actuation-workspace unit. Special care was taken to manually inspect the final design space results and ensure that they were not skewed toward hands with small reachable workspaces (small Q values), because very high values of H can be obtained when Q is small. The overall score for a hand design is the average score H_O over all simulated objects, as shown in Fig. 2 (written simply as H, as in the color bar label).

Kinematic topology enumeration

Using the previously described manipulation metrics, this work compares hand designs consisting of two opposing symmetric fingers, each modeled as an open serial chain linkage. In this section, we enumerate all possible planar serial chain finger topologies consisting of up to three revolute and prismatic joints. We leave hands with more than three joints per finger as future work, because it is computationally quite intensive to consider all designs of all permutations of kinematic topologies exhaustively with the current simulation.

We begin by limiting the scope to open serial chains, meaning that we are interested in mechanisms where one end is fixed and the opposite end is free, and all joints are in series with one another. Although it is likely that fingers based on parallel mechanisms would create highly dexterous hands, we choose to save their consideration for future work, allowing us to consider a tractable space of kinematic topologies herein. Specifically, we focus on planar

open serial chain linkages composed of revolute (R) and prismatic (P) joints, separated by rigid links, where each linkage has a fixed base link (the palm). The number of possible finger topologies for a finger with n actuators is 2^n and is found by counting the number of ordered samples with replacement for two actuation styles (R and P). The full list of finger topologies considered in this work is R, P, RR, PP, RP, PR, RRR, PPP, RRP, PPR, RPR, PRP, RPP, and PRR.

Design space variation

To determine how design can be leveraged for better manipulation, 6250 unique hand designs were simulated manipulating 10 distinct objects over their entire workspace, and the manipulation ability of each hand-object system was graded on the basis of the metric described by Eq. 10. All topologies enumerated in the previous section were simulated.

Key design parameters were varied to exhaustively explore the design space of planar open serial chain fingers. Specifically, the palm width p (the spacing between the fingers), the distal-most link length d, and the angle of each prismatic joint with respect to the previous link ϕ (typically the θ parameter in Denavit-Hartenberg notation) were all varied, as shown in Fig. 7. Each finger length (including half of the palm) was set to a dimensionless value of 1, so the values of these parameters can be thought of as proportions of total finger length rather than absolute values, permitting straightforward scaling of the system according to expected object sizes. In general, the finger link lengths were determined as follows. First, half of the palm width and the distal-most link length were subtracted from the overall finger length of 1. Next, the remaining portion of the finger was divided into $n - 1$ equal parts, where n is the number of actuators in the finger. Some special cases exist; for example, in the case of topologies containing only one actuator, the distal link length is set to zero, and the length of the single link is $1 - p$. In addition, when the first joint in a finger is prismatic, the length of its link is set equal to half the palm width—in other words, a proximally located prismatic link is considered to be part of the palm. Mathematically, the assignment of link lengths can be expressed as

$$l = \begin{cases} 1 - p, n = 1 \\ 1 - p - d, n > 1 \end{cases} \quad (11)$$

where p is half of the palm width, d is the distal link length, and l is the length of all other links. In practice, a palm width and a distal link length are prescribed, and the remaining link lengths are solved for using Eq. 11. The parameter p was varied from 0 to 0.5 (a proportion of overall finger length) in five steps, and the parameter d was varied from 0.0625 to 0.4 in five steps. As stated, the fixed angle of prismatic joints was varied from 0 to $\pi/2$, as shown in Fig. 2. This is equivalent to varying the Denavit-Hartenberg parameter normally represented by θ, which is the controllable parameter of a revolute joint but remains fixed for a prismatic joint. In addition, note that the transmission ratio K is different for R and P actuators. This is due to the gearing required to produce the linearly actuated prismatic joints, using standard rotary motors. In this case, we assume that prismatic joints are realized with a rack and pinion and that the pinion radius is 0.15 (again a proportion of overall finger length), but other values can be considered depending on design constraints. This value was chosen to approximately adjust each pinion motor's backdrivable range with that of the other motors located at revolute joints (so the pinion motor has about $\pi/2$ rad of usable range while running the carriage over the desired distance on the linear rail) to equalize the energy cost of displacing both P and R joints.

Object size, shape, and pose variation

Each hand was simulated while varying the pose, shape, and size of a target object. Specifically, each hand was simulated manipulating 10 distinct objects—5 circles and 5 squares. Each shape's radius was varied over five sizes, from 0.15 to 0.4 (a proportion of the total finger length). The pose of each object was varied to determine manipulability over the hand's entire workspace. Because the system is planar, the pose consists of an x coordinate and a y coordinate of the object's center and an orientation β. For circular objects, orientation was not varied. The metric detailed in Eq. 10 was computed at each valid pose for each object.

A valid pose of the object is one that is reachable by both fingers of the hand. If contact cannot be made between the hand and the object in a particular pose, the metric was not evaluated, because that pose is not within the hand-object system's reachable workspace. The range of simulated object sizes is shown in Fig. 7. Although the y coordinate of the object pose was allowed to extend to 0, in practice, those y-coordinate values smaller than the object radius were excluded from the grid, because their inclusion would place the object in intersection with the palm. The grid sampled in the simulation is also shown in Fig. 7.

Simulated actuation

Joint limits were set for each joint, the range between these limits was discretized into a set of individual set points for each joint, and the Cartesian product of all sets was calculated to create the set of all possible combinations of actuation inputs for a particular finger. The actuation limits for each joint were chosen on the basis of typical limits seen in robotic hands. For instance, the first revolute joint in a hand has a range between 0 and π rad (where both fingers at 0 corresponds to the hand being fully open), allowing it to sweep from one side of the palm to the other. Successive revolute joints, however, were limited to a range between 0 and $\pi/2$ rad with respect to the previous link to prevent excessive collision with other parts of the hand or hyperextension. Prismatic joint ranges were also selected on the basis of practical considerations, such as the approximate doubling in length that can be achieved with standard leadscrew-driven linear actuators. Thus, their actuation was limited between $l/2$ and l. However, prismatic joints located proximally in the finger were allowed to extend from 0 to l because clever packaging is possible within the palm. Special care was taken to ensure that this treatment did not skew the results toward kinematic topologies with proximally located prismatic joints, by testing with (0 to l) and without ($l/2$ to l) this condition and by ensuring that results were normalized by workspace size and did not tend toward hand designs with large proximal prismatic links.

SUPPLEMENTARY MATERIALS

robotics.sciencemag.org/cgi/content/full/6/54/eabd2666/DC1

Fig. S1. Open-loop operation through grasp primitive–based control.

Fig. S2. Object teleoperation through grasp primitive–based control.

Fig. S3. Visual servoing of object pose through grasp primitive–based control.

REFERENCES AND NOTES

1. G. Vassura, A. Bicchi, Whole-hand manipulation: Design of an articulated hand exploiting all its parts to increase dexterity, in *Robots and Biological Systems: Towards a New Bionics?* (Springer, 1993), pp. 165–177.
2. K. Salisbury, B. Roth, Kinematic and force analysis of articulated mechanical hands. *J. Mech. Trans. Autom. Design* **105**, 35–41 (1983).
3. J. Kerr, B. Roth, Analysis of multifingered hands. *Int. J. Robot. Res.* **4**, 3–17 (1986).
4. D. L. Brock, Enhancing the dexterity of a robot hand using controlled slip, in *Proceedings of the 1988 IEEE International Conference on Robotics and Automation* (IEEE, 1988), pp. 249–251.
5. J. C. Trinkle, A quasi-static analysis of dextrous manipulation with sliding and rolling contacts, in *Proceedings of the 1989 International Conference on Robotics and Automation* (IEEE, 1989), pp. 788–793.
6. R. D. Howe, M. R. Cutkosky, Practical force-motion models for sliding manipulation. *Int. J. Robot. Res.* **15**, 557–572 (1996).
7. K. Yu, M. Bauza, N. Fazeli, A. Rodriguez, More than a million ways to be pushed. A high-fidelity experimental dataset of planar pushing, in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2016), pp. 30–37.
8. T. Okada, Computer control of multijointed finger system for precise object-handling. *IEEE Trans. Syst. Man Cybern.* **12**, 289–299 (1982).
9. A. Cole, J. Hauser, S. Sastry, Kinematics and control of multifingered hands with rolling contact, in *Proceedings of the 1988 IEEE International Conference on Robotics and Automation* (IEEE, 1988), pp. 228–233.
10. D. J. Montana, The kinematics of contact and grasp. *Int. J. Robot. Res.* **7**, 17–32 (1988).
11. W. Yuan, S. Dong, E. H. Adelson, GelSight: High-resolution robot tactile sensors for estimating geometry and force. *Sensors* **17**, 2762 (2017).
12. A. Bicchi, Hands for dexterous manipulation and robust grasping: A difficult road toward simplicity. *IEEE Trans. Robot. Autom.* **16**, 652–662 (2000).
13. R. R. Ma, A. M. Dollar, On dexterity and dexterous manipulation, in *2011 15th International Conference on Advanced Robotics (ICAR)* (IEEE, 2011).
14. A. M. Okamura, N. Smaby, M. R. Cutkosky, An overview of dexterous manipulation, in *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065)* (IEEE, 2000), vol. 1.
15. S. Makita, Y. Maeda, 3D multifingered caging: Basic formulation and planning, in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, Nice, France, 2008), pp. 2697–2702.
16. R. Diankov, S. S. Srinivasa, D. Ferguson, J. Kuffner, Manipulation planning with caging grasps, in *Humanoids 2008-8th IEEE-RAS International Conference on Humanoid Robots* (IEEE, 2008), pp. 285–292.
17. A. Rodriguez, M. T. Mason, S. Ferry, From caging to grasping. *Int. J. Robot. Res.* **31**, 886–900 (2012).
18. L. Birglen, C. M. Gosselin, Geometric design of three-phalanx underactuated fingers. *J. Mech. Des.* **128**, 357–364 (2006).
19. A. M. Dollar, R. D. Howe, The highly adaptive SDM hand: Design and performance evaluation. *Int. J. Robot. Res.* **29**, 585–597 (2010).
20. F. L. Hammond, J. Weisz, A. Andrés, P. K. Allen, R. D. Howe, Towards a design optimization method for reducing the mechanical complexity of underactuated robotic hands, in *2012 IEEE International Conference on Robotics and Automation* (IEEE, 2012), pp. 2843–2850.
21. M. Ciocarlie, F. M. Hicks, R. Holmberg, J. Hawke, M. Schlicht, J. Gee, S. Stanford, R. Bahadur, The Velo gripper: A versatile single-actuator design for enveloping, parallel and fingertip grasps. *Int. J. Robot. Res.* **33**, 753–767 (2014).
22. R. R. Ma, N. Rojas, A. M. Dollar, Spherical hands: Toward underactuated, in-hand manipulation invariant to object size and grasp location. *J. Mech. Robot.* **8**, 061021 (2016).
23. W. G. Bircher, A. M. Dollar, N. Rojas, A two-fingered robot gripper with large object reorientation range, in *2017 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2017), pp. 3453–3460.
24. A. Kakogawa, H. Nishimura, S. Ma, Underactuated modular finger with pull-in mechanism for a robotic gripper, in *2016 IEEE International Conference on Robotics and Biomimetics (ROBIO)* (IEEE, 2016), pp. 556–561.
25. C. M. McCann, A. M. Dollar, Design of a Stewart platform-inspired dexterous hand for 6-DOF within-hand manipulation, in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2017), pp. 1158–1163.
26. V. Tincani, M. G. Catalano, E. Farnioli, M. Garabini, G. Grioli, G. Fantoni, A. Bicchi, Velvet fingers: A dexterous gripper with active surfaces, in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2012), pp. 1257–1263.
27. M. Ceccarelli, C. Lanni, A multi-objective optimum design of general 3R manipulators for prescribed workspace limits. *Mech. Mach. Theory* **39**, 119–132 (2004).
28. C. Hazard, N. Pollard, S. Coros, Automated design of robotic hands for in-hand manipulation tasks. *Int. J. Hum. Robot.* **17**, 1950029 (2020).
29. A. Meixner, C. Hazard, N. Pollard, Automated design of simple and robust manipulators for dexterous in-hand manipulation tasks using evolutionary strategies, in *2019 IEEE-RAS 19th International Conference on Humanoid Robots (Humanoids)* (IEEE, 2019), pp. 281–288.
30. F. Sanfilippo, G. Salvietti, H. X. Zhang, H. P. Hildre, D. Prattichizzo, Efficient modular grasping: an iterative approach, in *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)* (IEEE, 2012), pp. 1281–1286.
31. R. R. Ma, A. M. Dollar, Linkage-based analysis and optimization of an underactuated planar manipulator for in-hand manipulation. *J. Mech. Robot.* **6**, 011002 (2014).
32. J. Borràs, A. M. Dollar, Dimensional synthesis of three-fingered robot hands for maximal precision manipulation workspace. *Int. J. Robot. Res.* **34**, 1731–1746 (2015).
33. L. U. Odhner, A. M. Dollar, Stable, open-loop precision manipulation with underactuated hands. *Int. J. Robot. Res.* **34**, 1347–1360 (2015).
34. B. Sundaralingam, T. Hermans, Relaxed-rigidity constraints: In-grasp manipulation using purely kinematic trajectory optimization. *Planning* **6**, 7 (2017).
35. Y. Maeda, T. Asamura, Sensorless in-hand caging manipulation, in *International Conference on Intelligent Autonomous Systems* (Springer, 2016), pp. 255–267.
36. J. C. Trinkle, R. Ram, A. Farahat, P. F. Stiller, Dexterous manipulation planning and execution of an enveloped slippery workpiece, in *1993 Proceedings IEEE International Conference on Robotics and Automation* (IEEE, 1993), pp. 442–448.
37. Z. Li, P. Hsu, S. Sastry, Grasping and coordinated manipulation by a multifingered robot hand. *Int. J. Robot. Res.* **8**, 33–50 (1989).
38. R. R. Ma, W. G. Bircher, A. M. Dollar, Toward robust, whole-hand caging manipulation with underactuated hands, in *2017 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2017), pp. 1336–1342.
39. R. R. Ma, W. G. Bircher, A. M. Dollar, Modeling and evaluation of robust whole-hand caging manipulation. *IEEE Trans. Robot.* **35**, 549–563 (2019).
40. W. G. Bircher, A. S. Morgan, K. Hang, A. M. Dollar, Energy gradient-based graphs for planning within-hand caging manipulation, in *2019 International Conference on Robotics and Automation (ICRA)* (IEEE, 2019), pp. 2462–2467.
41. G. A. Kragten, J. L. Herder, The ability of underactuated hands to grasp and hold objects. *Mech. Mach. Theory* **45**, 408–425 (2010).
42. M. T. Mason, A. Rodriguez, S. S. Srinivasa, A. S. Vazquez, Autonomous manipulation with a general-purpose simple hand. *Int. J. Robot. Res.* **31**, 688–703 (2012).
43. W. G. Bircher, A. M. Dollar, Design principles and optimization of a planar underactuated hand for caging grasps, in *2019 International Conference on Robotics and Automation (ICRA)* (IEEE, 2019), pp. 1608–1613.
44. F. J. Romero-Ramirez, R. Muñoz-Salinas, R. Medina-Carnicer, Speeded up detection of squared fiducial markers. *Image Vis. Comput.* **76**, 38–47 (2018).
45. B. Calli, A. Singh, A. Walsman, S. Srinivasa, P. Abbeel, A. M. Dollar, The YCB object and model set: Towards common benchmarks for manipulation research, in *2015 International Conference on Advanced Robotics (ICAR)* (IEEE, 2015), pp. 510–517.
46. A. J. Spiers, B. Calli, A. M. Dollar, Variable-friction finger surfaces to enable within-hand manipulation via gripping and sliding. *IEEE Robot. Autom. Lett.* **3**, 4116–4123 (2018).
47. S. Boyd, L. Vandenberghe, *Convex optimization* (Cambridge Univ. Press, 2004).
48. G. M. Ziegler, *Lectures on Polytopes* (Springer Science & Business Media, 2012), vol. 152.

Acknowledgments: W.G.B. would like to thank S. Muchnik for making the soft knitted cube and for posing as a hand model for Fig. 3 and C. Harshaw, E. Muchnik, and A. Yawar for discussions related to mathematical notation. **Funding:** This research was supported by the U.S. NSF under grants IIS-1734190 and IIS-1900681. **Author contributions:** W.G.B. contributed all theoretical developments and carried out the design space search, the hand design and assembly, the controller design, and the test object experiments. A.S.M. was responsible for all experiments using the WAM robotic arm and the YCB Object Set, as well as for making the video. A.M.D. provided direction and insight during the duration of the project, as well as grant funding and a laboratory space and resources. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions of this paper are available in the paper or the Supplementary Materials.

Submitted 2 August 2020
Accepted 19 April 2021
Published 12 May 2021
10.1126/scirobotics.abd2666

Citation: W. G. Bircher, A. S. Morgan, A. M. Dollar, Complex manipulation with a simple robotic hand through contact breaking and caging. *Sci. Robot.* **6**, eabd2666 (2021).

Science Robotics

Complex manipulation with a simple robotic hand through contact breaking and caging

Walter G. Bircher, Andrew S. Morgan, and Aaron M. Dollar

Sci. Robot. **6** (54), eabd2666. DOI: 10.1126/scirobotics.abd2666

Science Robotics (ISSN 2470-9476) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Robotics* is a registered trademark of AAAS.