Comparative Peg-in-Hole Testing of a Force-Based Manipulation Controlled Robotic Hand

Karl Van Wyk, Mark Culleton, Joe Falco, and Kevin Kelly

***Abstract*—Force-based manipulation control strategies are evolving as a primary mechanism in robotics for performing the fine manipulation tasks typical within manufacturing assembly. The ability to systematically compare robotic system performance and quantify true advancement in fine manipulation is of utmost importance. Accordingly, the objectives of this paper are threefold: 1) creation of a peg-in-hole test method with associated performance metrics and a systematic data analysis strategy for performance benchmarking, 2) first demonstration of a recently developed manipulation controller piloting a robotic hand and its paired task-level logic for completing the peg-in-hole test, and 3) exemplifying the performance benchmarking technique by comparing two approaches for robotic insertions—the previously mentioned compliant hand, stiff arm system, and a stiff gripper, compliant arm system. Analyses reveal that the unconventional hand system can perform at and sometimes above the level of the gripper system in the developed peg-in-hole scenario. Moreover, the hand's active control of the peg's full Cartesian pose reduces positional error sensitivity and minimizes exerted insertion forces, highlighting the strategy's potential for fine manipulation tasks. Results indicate that robotic arms equipped with highly articulated and sensorized robotic hands can provide a truly realizable solution path for performing peg-in-hole tasks.**

***Index Terms*—Grasping, manipulation, performance testing.**

I. Introduction

The ability to reliably perform part insertions, a relatively simple task for humans, remains a challenging problem for robotic systems. Two components are necessary to overcome this problem, including a wide adoption of benchmarks that objectively compare and contrast findings by active researchers in the field, and the development of innovative solutions in robotic manipulation.

One necessary component for performance benchmarking is a physical test with predefined artifacts, parameters, rules, initial conditions, and a data analysis procedure. Similar to other robotic benchmarks [1], [2] and robotic hand tests from the National Institute of Standards and Technology [3], this research effort created an easily replicable physical test with data analytics to help gauge the performance of robotic systems for insertion operations. This test is based on a peg-in-hole task, a particularly relevant task for industrial robots as it accounts for over 35 % of all assembly operations [4]. Others have also created peg-in-hole benchmarking tests in the field of teleoperation [5]–[9], and human rehabilitation and prosthetics [10]–[12]. These existing benchmarking tests possess one or some combination of the following shortcomings:

1) arbitrary design of test artifacts,
2) completion time (CT) as sole performance measure,
3) prohibitive cost of designed tests through either (or both) the use of force measurement or custom machined parts,
4) incomplete suite of relevant statistical algorithms that can be applied to both continuous and discontinuous types of data, and
5) excluding preventative measures for correlated performance samples.

In contrast, the peg-in-hole test highlighted herein was designed with reference to data acquired from humans in performing assembly operations [13]. The artifacts are designed for low-cost replication through three-dimensional (3-D) printing. Both CT and probability of success (PS) are reported as important and inexpensively obtainable performance measures. Finally, a carefully selected set of nonparametric statistical algorithms is suggested for analyzing the performance data, while the underlying assumption of sample independence for valid analyses is guaranteed through applying randomly generated hole position errors during testing.

In terms of bestowing robotic systems with the fine motor skills required for insertion processes, one promising avenue includes sophisticated sensing and control of robotic hands. At its core, multifingered grasping and manipulation operations can be mathematically cast by seeking to control the Cartesian pose of an object in the environment. Typical approaches to this problem are either kinematic [14], [15] or kinetic [16], [17] in nature. In order to test the theory that complex robotic hand technology is approaching readiness levels for application to real-world problems, this work presents experimental results on a recently developed force-based manipulation controller [18].

Two main robotic systems were assessed using the developed test method. One system used a compliant robotic hand with the aforementioned manipulation control strategy. The other system leveraged more standard solutions to peg-in-hole that involve searching with a rigid end-effector coupled to a compliant or force-controlled arm [19]. Since both approaches afford system compliance for insertion operations, the case study herein seeks to identify and quantify statistically significant performance differences between the two viable strategies. Systematic statistical analyses reveal that the approach leveraging a force-controlled robotic hand can compete with standard approaches to peg-in-hole insertion and, in some circumstances, can exceed the performance of those approaches.

In order to methodically cover the important aspects of this research, the paper is organized as follows. Section II discusses the peg-in-hole

Manuscript received July 5, 2017; revised November 17, 2017; accepted December 7, 2017. Date of publication February 19, 2018; date of current version April 12, 2018. This paper was recommended for publication by Associate Editor A. Franchi and Editor A. Kheddar upon evaluation of the reviewers' comments. This work was supported in part by the Intelligent Systems Division at the National Institute of Standards and Technology and in part by the Irish Research Council under the "Embark Initiative." *Disclaimer:* Certain commercial equipment, instruments, or materials are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose. *(Corresponding author: Karl Van Wyk.)*

K. Van Wyk and J. Falco are with the National Institute of Standards and Technology, Gaithersburg, MD 20899 USA (e-mail: karl.vanwyk@nist.gov; joseph.falco@nist.gov).

M. Culleton and K. Kelly are with the Department of Mechanical and Manufacturing Engineering, Trinity College Dublin, Dublin 2, Ireland (e-mail: culletom@tcd.ie; kevin.kelly@tcd.ie).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TRO.2018.2791591

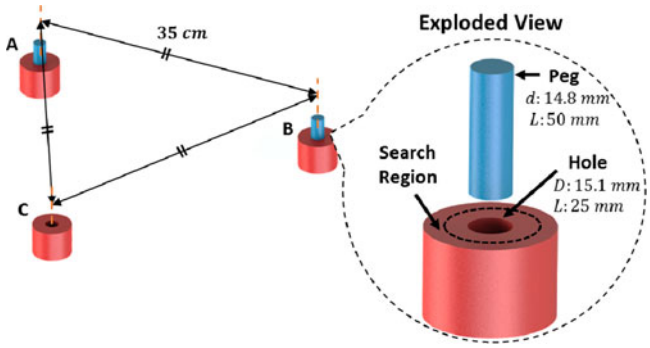

Fig. 1. Peg-in-hole test setup and design specifications.

test method, metrics, and data analysis routines, Section III discusses the robotic systems and strategies used to solve the peg-in-hole test, Section IV gives an in-depth analysis of system performance, and Section V offers conclusive statements derived from the data analyses along with final remarks regarding the robotic systems.

II. Peg-in-Hole Test Method and Performance Analysis Strategy

A. Design and Setup

To represent a typical peg-in-hole assembly task, the test is set up, as shown in Fig. 1. The holes are spaced approximately 35 cm apart, which is the upper threshold for primary assembly processes (i.e., operations that directly contribute to the formation of a product) [20]. The pegs and holes were designed to be 3-D printable to reduce manufacturing costs and promote test replication.[1] Acrylonitrile butadiene styrene (ABS) filament is recommended for its larger impact strength. The peg diameter d and hole diameter D are chosen to give a dimensionless clearance $c = (D - d)/D = 0.02$. This clearance defines a sufficiently difficult insertion task, as it is small enough to influence a human assembly worker's insertion time [13].

As a reference point, the classification tables from the Boothroyd–Dewhurst design-for-assembly (DFA) method [13] can be used to estimate human-level performance. DFA methods are used to reduce the overall cost of an assembly by minimizing the number of discrete operations, the number of parts, and the complexity of the remaining operations. As part of its optimization approach, the Boothroyd–Dewhurst DFA method has developed classification tables that identify the fundamental operations within an assembly and quantify their difficulty using human CTs. These CTs represent the average time taken by a human to perform each operation and are based on empirical data that were collected over a period of years [21]. The Boothroyd–Dewhurst DFA method is the most widely used and accepted DFA method [22], as its consideration of both handling factors (e.g., part size, weight, symmetry, and properties) and insertion factors (e.g., insertion symmetry, ease of reach, insertion resistance, alignment, positioning, holding, and fastening) enables the accurate estimation of overall assembly time. Accordingly, the Boothroyd–Dewhurst classification tables provide a means to reasonably estimate human performance without the need for task-specific testing and analysis. In this scenario, the classification tables identify the peg insertion task as a single "01" insertion operation, which yields an estimated total CT of 2.5 s.

[1]CAD models for the test artifacts are freely available at https://www.nist.gov/programs-projects/performance-metrics-and-benchmarks-advance-state-robotic-assembly

The test initializes with two pegs placed into two holes. The goal is to cyclically transfer a peg into the next available hole. The planar hole location prior to every insertion attempt is intentionally misaligned by adding positional offsets in both the *X*- and *Y*-directions. These offsets represent positional uncertainty (from control, calibration, or perception error) and were drawn from two Gaussian distributions: 1) zero mean and 1 mm standard deviation (σ_1), and 2) zero mean and 2 mm standard deviation (σ_2). The test was conducted with 60 attempted insertions, where the *X*–*Y* positional error was duplicated across both systems for each insertion attempt.

The application of positional error is an important aspect of this test method for multiple reasons. First, its presence introduces stochasticity into robot's performance data (many statistical tests assume sufficiently independent, uncorrelated samples). Furthermore, controlling the positional error allows for performance benchmarking by easily subjecting different robotic systems to the same positional errors, or subjecting one robotic system to various levels of positional error.[2]

B. Performance Measures

Two systemic metrics have been employed to capture the performance of a robotic system during this task-level assessment.

1) Completion Time (CT): The time required to successfully complete a peg insertion. This metric incorporates the time required to align and insert the peg. Peg acquisition and transportation times were constant throughout testing and are not included in the presented CT. This prevents the CT being subjugated by robot arm motions and provides a truer reflection of the performance of each insertion strategy.

2) Probability of Success (PS): The probability of a successful peg insertion. The PS value is calculated by relating the number of samples, the number of successes, and a confidence level [23]. This measure employed a 95 % confidence level, a standard setting used within the industry [24].

C. Comparative Statistics

The performance measures for this test involve the statistical analysis of the measured CT and PS across all systems.[3]

1) Continuous Data: For the CT data, an autocorrelation coefficient of lag one (r_1) is calculated to verify that the collected samples are statistically independent, an underlying assumption for the subsequent tests. Next, a Kolmogorov–Smirnov (KS) test is conducted to determine whether or not the distributions of two CT datasets are significantly different. Distribution tests are useful first-line indicators for detecting differences in datasets. They analyze the overall shape of the datasets instead of just the first or second statistical moments. Following, a Levene test with the Brown–Forsythe statistic is used to analyze the variances (σ^2) of two CT datasets (ANOVA). This particular ANOVA formulation generalizes well as it provides robustness against non-normal data [25]. Depending on the outcome of the ANOVA test, the appropriate variant of the *t*-test is applied to determine if there is a difference in the sample means (μ) of two CT datasets.

2) Discontinuous Data: Since the data for calculating PS values are discrete by nature (number of successes and failures), the Kolmogorov–Conover (KC) algorithm [26] (designed for analyzing ordinal datasets) is used to determine if a statistical difference exists

[2]The applied positional error and returned robotic performance data can be downloaded from https://www.nist.gov/programs-projects/performance-metrics-and-benchmarks-advance-state-robotic-assembly

[3]Software for statistical tests is freely available at https://www.nist.gov/el/intelligent-systems-division-73500/performance-data-analytics

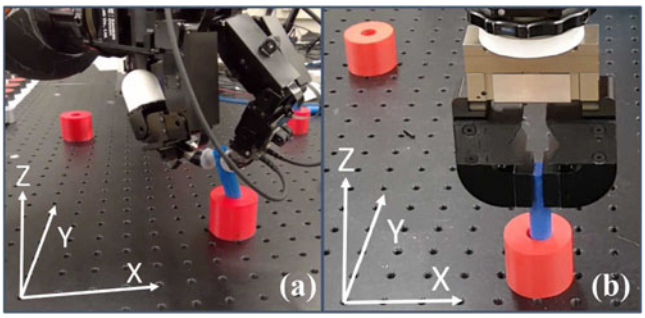

Fig. 2. Attempted peg insertion by (a) system 1 and (b) system 2.

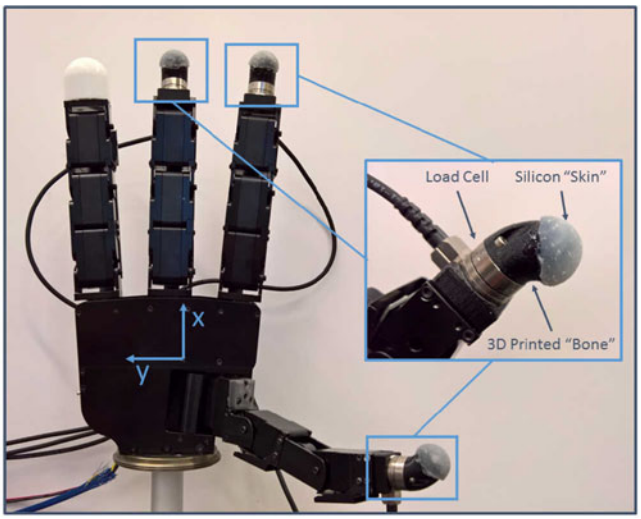

Fig. 3. Robotic hand of system 1.

between two sets of pass–fail data. The statistical tests were conducted across the datasets at a 95 % confidence level.

III. Robotic Systems and Solutions

In general, robots should employ some formulation of compliance or force control when making contact with their environment. Therefore, two distinct force control strategies were compared using the defined performance benchmarking technique: 1) a 7 degree-of-freedom, Cartesian position-controlled arm with an in-hand object Cartesian impedance-controlled robotic hand (system 1); and 2) a 7 degree-of-freedom, Cartesian impedance-controlled arm with a pneumatic, parallel gripper (system 2), as shown in Fig. 2. System 1 achieves compliance through the robotic hand, while system 2 achieves compliance through the robotic arm. The same robotic arm was used, but with a different position controller.

A. System 1

1) Hand Mechanics and Sensory Suite: The robotic hand has four fingers, 16 independently actuated joints with rotary encoders, and three 6-axis force–torque transducers at the fingertips (see Fig. 3). The sensing suite includes a touch-based, 6-DoF object pose estimation algorithm, 3-D fingertip force, 3-D fingertip normal force, and 3-D fingertip center of pressure. For brevity, the custom algorithms providing these sensing modalities are not discussed here. All sensing and control rates operated at a nominal 333 Hz.

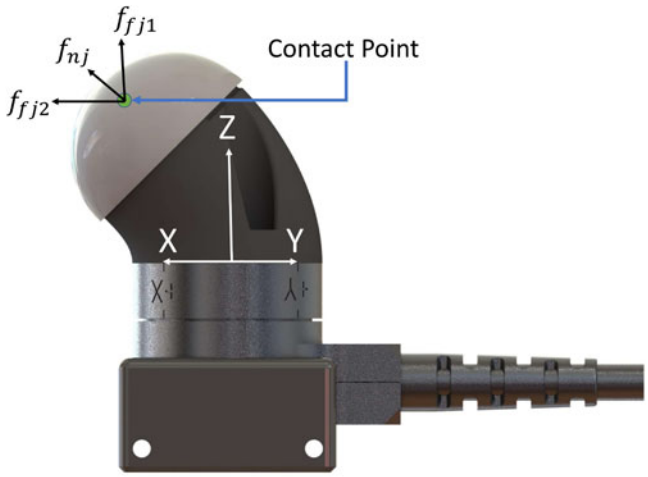

Fig. 4. Load cell coordinate system and illustration of normal and friction force basis vectors.

2) Multifingered Manipulation Controller: A key feature for system 1 is the hand's ability to perform in-hand peg manipulation, not just grasping. Much like the intuition behind Cartesian control for existing robotic manipulators, the manipulation controller for the robotic hand was sculpted from the desire to control the Cartesian pose of an object. The implemented controller is a slightly modified version of the one developed in [18] and essentially enables impedance control of the object via finger Cartesian force control. Given a desired and measured Cartesian position profile for the object $r_{\mathrm{cd}} \in \mathbb{R}^{6\times 1}$ and $r_c \in \mathbb{R}^{6\times 1}$, respectively, the tracking error signal is

$$e \triangleq r_{\mathrm{cd}} - r_c. \tag{1}$$

The manipulation controller calculates the desired Cartesian force profile $f_{\mathrm{jd}} \in \mathbb{R}^{3\times 1}$ for every finger j, $j = 1, ..., k$ ($k = 3$ since only three of the four fingers were used) from

$$f_{\mathrm{jd}} = \hat{f}_{\mathrm{nj}} u_{\mathrm{nj}} + \hat{f}_{\mathrm{fj1}} u_{\mathrm{fj1}} + \hat{f}_{\mathrm{fj2}} u_{\mathrm{fj2}} \tag{2}$$

where $\hat{f}_{\mathrm{nj}} \in \mathbb{R}^{3\times 1}$ is the jth unitized contact normal force, and $\hat{f}_{\mathrm{fj1}} \in \mathbb{R}^{3\times 1}$ and $\hat{f}_{\mathrm{fj2}} \in \mathbb{R}^{3\times 1}$ are the jth measured friction force basis vectors. The force–torque transducers were calibrated to resolve 3-D contact points on the surface of the fingertip. There are many strategies for performing this calibration that can be analytical [27] or numerical in nature. Coupling the knowledge of contact location with the known geometry of the fingertip, one can readily calculate the directionality of the normal force at the point of contact. There are several options for calculating friction force basis vectors, but in this case, $\hat{f}_{\mathrm{fj1}} = \hat{f}_{\mathrm{nj}} \times Z$ and $\hat{f}_{\mathrm{fj2}} = \hat{f}_{\mathrm{nj}} \times \hat{f}_{\mathrm{fj1}}$, where Z is the Z-axis of the sensor's coordinate system (see Fig. 4). Furthermore, u_{nj}, u_{fj1}, and $u_{\mathrm{fj2}} \in \mathbb{R}$ are the signed force magnitudes in their respective directions, as indicated in (2). Control laws are explicitly established for $U_n \in \mathbb{R}^{k\times 1}$ and $U_f \in \mathbb{R}^{2k\times 1}$ where

$$U_n = [u_{\mathrm{n1}}, ..., u_{\mathrm{nk}}]^T \tag{3}$$

$$U_f = [u_{\mathrm{f11}}, u_{\mathrm{f12}} ..., u_{\mathrm{fk1}}, u_{\mathrm{fk2}}]^T. \tag{4}$$

The control laws are

$$U_n = \gamma_1 \mathrm{Sig}\left(B_n^T (k_p e + k_d \dot{e}) - \gamma_2 I_{k1}\right) + \gamma_3 I_{k1} \tag{5}$$

$$U_f = B_f^T (k_p e + k_d \dot{e}) \tag{6}$$

where $B_n \in \mathbb{R}^{6\times k}$ and $B_f \in \mathbb{R}^{6\times 2k}$ are the normal and friction grasp matrices, γ_1, γ_2, and γ_3 are positive control shaping scalars, k_p and

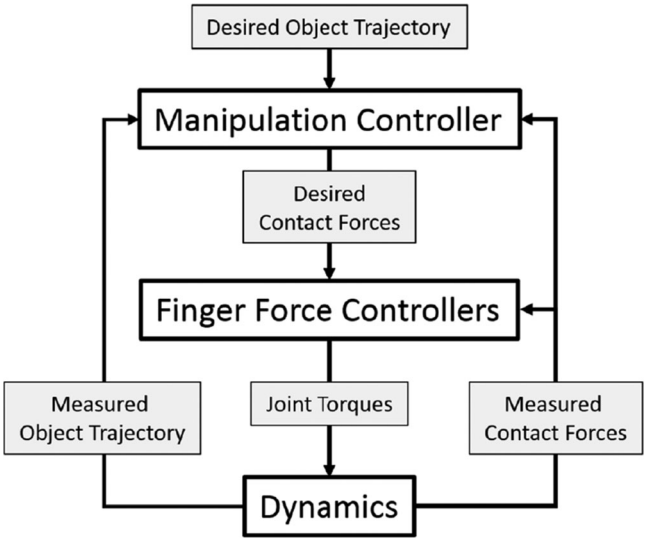

Fig. 5. Multifingered manipulation control architecture.

k_d are the stiffness and damping parameters, Sig(.) is the sigmoid function of vector elements, and $I_{k1} \in \mathbb{R}^{k\times 1}$ is a vector of ones. For this application, the control shaping scalars were set as $\gamma_1 = 4$, $\gamma_2 = 5$, $\gamma_3 = 0.35$. This ensures that each finger emits at least 0.35 N of normal force and at most 4.35 N of normal force on the object, the minimum and maximum forces any one finger could reliably control based on finger strength and force sensing. The control gains were set as $k_p = 0.4$ N/mm and $k_d = 0.02$ N·s/mm for translational motion, and $k_p = 0.2$ N/mm·rad and $k_d = 0.02$ N·s/mm·rad for rotational motion. These values were hand-tuned for stable grasping and manipulation operation of the peg by the hand (although regions of stiffness and damping values for stable control exist). Henceforth, the performance of this system was only reported with these preset parameter values. Note, (6) has an accompanying projection algorithm to prevent issuing friction forces outside the contact friction cone. In this particular case, the coefficient of static friction μ_s was set to 0.3, a reasonable value for rubber and plastic (materials of fingertip and peg). Furthermore, r_c is continuously estimated by the hand's forward kinematics. Refer to [18] for more information on this control scheme.

Once f_{jd} is calculated for every finger, these commands are issued to an underlying Cartesian force controller that operates for each finger, individually (see Fig. 5). For brevity, details regarding the finger force controller are not discussed here.

3) Task-Level Control Strategy: System 1 conducted the test with the hand actively controlling the pose of the peg using the aforementioned sensing and control technology while coordinating motions with the arm (see Fig. 6). All motions are commanded and force measurements made with respect to the hand's coordinate system (see Figs. 3 and 6 for coordinate system placement). The insertion strategy is enumerated as follows.

1) Hand tilts the peg 35° about the *Y*-axis to increase the likelihood of a collision between the edge of the peg and the edge of the hole during the descent [see Fig. 6(a)].
2) Arm translates along the positive *Z*-axis until a resultant force of greater than 0.1 N is measured by calculating the magnitude of the sum of the contact forces across the three fingertip load cells. Upon sufficient contact, if the magnitude of the resultant contact force is not greater than 0.085 N in either the *X*- or *Y*-axes, then randomly prod within a 4 mm × 4 mm search grid until this force requirement is met. This ensures that the peg is sufficiently engaged with the hole for further guidance [see Fig. 6(b) and (c)].
3) Arm translates by steps of −2 mm in the *X*-axis until the force in the *X*-direction is less than −0.06 N. Meanwhile, the arm also translates by 2 mm in the positive *Y*-axis if the resultant contact force in the *Y*-axis is less than −0.025 N and vice versa. Combined, these objectives ensure the tilted peg straddles the hole such that a subsequent negative rotation about the *Y*-axis and arm translation will partially insert the peg [see Fig. 6(c) and (d)].
4) Hand rotates the peg −35° about the *Y*-axis, while simultaneously, the arm translates the hand 15 mm in the *X*-axis and 4 mm in the *Z*-axis, which promotes a partial peg insertion state [see Fig. 6(e)].
5) Arm and hand each translate the peg 1 mm in the *Z*-axis until full insertion unless one of the following events occur. If the peg's

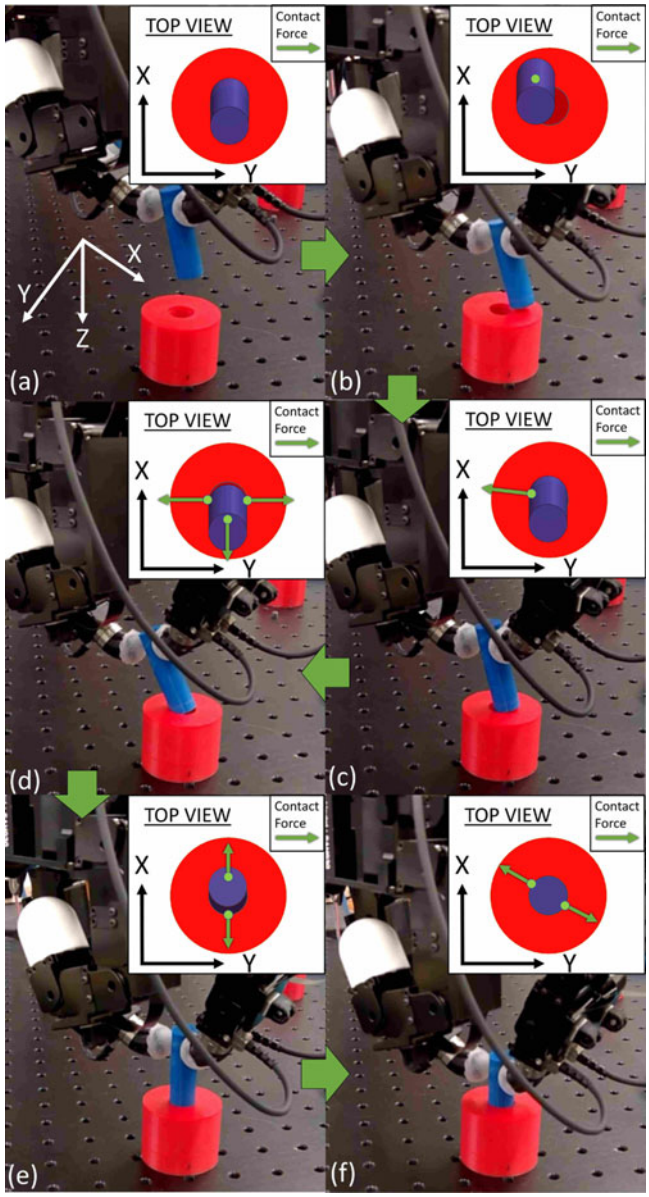

Fig. 6. System 1 task-level insertion process illustrating contact forces.

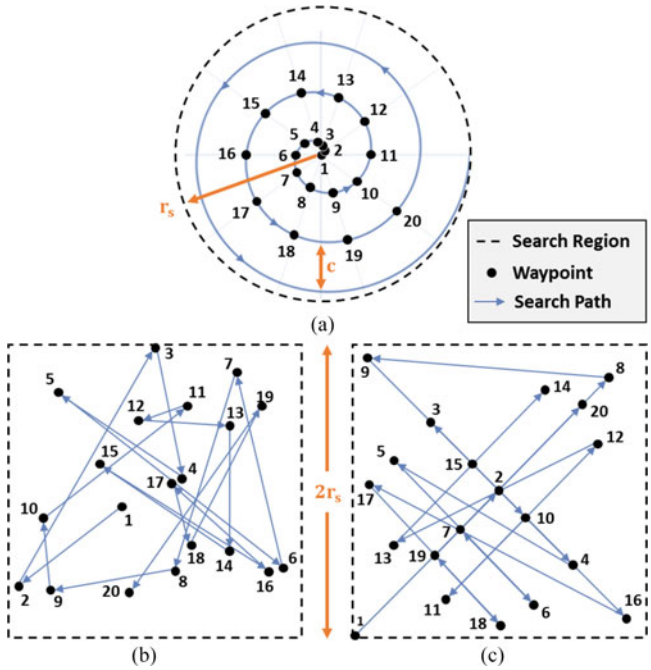

Fig. 7. Search routines used by system 2 showing first 20 waypoints: (a) Spiral search where turning distance equals the insertion clearance c (waypoints constant, turns enlarged for clarity), (b) random search (waypoints vary), and (c) quasi-random search (waypoints constant). Searches region radius r_s derived from the positional error.

rotation about the *X*-axis is greater than 5° or less than −5°, then the arm translates −1 or 1 mm in the *Y*-axis, respectively. Likewise, if the rotation about the *Y*-axis is greater than 5° or less than −5°, then the arm translates 1 or −1 mm in the *X*-axis. Furthermore, if the resultant contact force is positive in the *X*- or *Y*-axes, then the arm translates −1.5 mm in the respective axis. Similarly, a negative contact force in the *X*- or *Y*-axes results in a 1.5 mm translation in the respective axis. Exploiting both kinematic- and kinetic-based corrections promote continual peg–hole alignment during insertion while also minimizing contact forces [see Fig. 6(f)].

Note, this manipulation strategy for insertion is agnostic to the number of peg–hole contacts (illustrated in Fig. 6) since the insertion process relies solely on the resultant contact force and peg pose for guidance. In fact, the application of perception error and prodding by system 1 during the insertion process yielded a variety of contact scenarios between the peg and the hole. For instance, the state depicted in Fig. 6(d) yielded up to three points of contact.

B. System 2

1) Task-Level Control Strategy: System 2 conducts the peg-in-hole test by coupling the arm's impedance properties with one of three search routines: spiral, random, and quasi-random (see Fig. 7). During a search, the peg remains stationary within the gripper, while the robot arm is in Cartesian impedance control. The robot arm's stiffness and damping along the *Z*-axis is reduced, and the robot attempts to position the peg 10 mm below the surface of the block. If the peg and hole are not aligned, a contact force is generated and maintained during each search routine until the peg and hole are aligned. At this point, the robot's desired and measured positions match, and the control program identifies a successful insertion. The performance of system 2 was reported with a Cartesian stiffness of 2000 N/m in the *X*- and *Y*-axes (defaults), 1000 N/m in the *Z*-axis, and 200 N·m/rad about all axes (defaults).

2) Search Region: During peg insertion, a search region is defined based on the added positional error. To encompass the actual hole location, the radius of the search region r_s is set to 4σ, where σ is the standard deviation of the positional error. As the positional error follows a bivariate normal distribution with zero mean, this radius ensures that the hole will lie within the search region with a confidence level of 99.97 % [28]. Accordingly, r_s for positional error σ_1 and σ_2 is 4 and 8 mm, respectively. For the random and quasi-random search routines, a square search region with side length $2r_s$ is defined.

3) Spiral Search Routine: The spiral search is commonly used within the industry as it is an optimized search strategy for two-dimensional environment. It is particularly useful for Gaussian distributions, where the beginning of the spiral path coincides with the distributions mean [29]. When performing this search routine, system 2 follows an Archimedean spiral—a curve generated by moving away from a point at constant linear and angular velocity. In polar coordinates, the spiral is defined by

$$r = a\theta \tag{7}$$

where a is a real number and corresponds to the distance between successive turnings. The spiral search path is optimized for a peg-in-hole insertion operation by setting this distance to the insertion clearance c [30]. This minimizes the search path distance while still ensuring that the hole lies along the search path. The spiral is broken into ten motion waypoints per turn to accommodate command rate restrictions (see Fig. 7).

4) Random Search Routine: The random search is a particularly simple yet effective search strategy, as exemplified by its use by the majority of commercially available mobile robots [31]. When performing this search routine, system 2 attempts to insert the peg by moving the peg randomly within the defined search region. This is achieved by programming the robot to move to uncorrelated, randomly generated waypoints of uniform probability in both the *X*- and *Y*-directions.

5) Quasi-Random Search Routine: The quasi-random search routine utilizes a quasi-random sequence to generate the insertion waypoints. Stated simply, a quasi-random sequence fills a given space more uniformly than uncorrelated random points by subrandomly generating points that minimize the maximum distance between all points [32]. For this test, the points are generated using the two-dimensional Sobol sequence. This sequence is generated number-theoretically so that successive points fill the gaps in the previously generated distribution. The code for generating the Sobol sequence is presented and discussed in [32].

IV. Performance Comparisons

A. Cross-System Comparisons

Both systems performed the peg-in-hole test at the two defined levels of positional error, and numerical results are reported in Tables I and II. Analysis of performance data follows for all the aforementioned statistical tests. Comparisons are made between system 1 and all variants of system 2.

1) Autocorrelation, r_1: At 60 samples and 95% confidence level, an r_1 value beyond ± 0.25 indicates detectable correlation in the data. System 2 under the spiral search routine for σ_1 exhibited the largest r_1 value of 0.15, followed by system 2 (quasi-random) under σ_2 with an r_1 value of -0.11. This heightened correlation of the former is likely a result of the structured nature of the search routine coupled with low positional error. Despite this observation, all datasets are sufficiently

TABLE I
PERFORMANCE MEASURES OF ROBOTIC SYSTEMS ON PEG-IN-HOLE TASK WITH POSITIONAL ERROR σ_1

Robotic system	r_1	KS	μ (s)	σ^2 (s^2)	PS (%)
System 1	−0.07		11.70	40.57	85.4
System 2 *Spiral*	0.15	*	7.19*	12.79	95.2
System 2 *Random*	−0.04	*	8.01	59.92	95.2
System 2 *Quasi-Random*	−0.04	*	3.11*	8.84*	95.2

*Statistically significant difference when compared to system 1.

TABLE II
PERFORMANCE MEASURES OF ROBOTIC SYSTEMS ON PEG-IN-HOLE TASK WITH POSITIONAL ERROR σ_2

Robotic system	r_1	KS	μ (s)	σ^2 (s^2)	PS (%)
System 1	0.01		18.31	107.3	87.6
System 2 *Spiral*	0.07	*	37.13*	399.6*	95.2
System 2 *Random*	−0.01	*	15.62	417.72	95.2
System 2 *Quasi-Random*	−0.11	*	8.2*	50.25*	95.2

*Statistically significant difference when compared to system 1.

uncorrelated with low r_1 values, satisfying an underlying assumption for all subsequent statistical tests.

2) Distribution, KS: The distribution of performance data for system 1 was statistically different from that of system 2 across all search variants for σ_1 and σ_2 positional error levels. These preliminary indicators suggest that there is an increased chance—though not guaranteed—of detecting a difference in sample means or variances. For example, statistical differences were observed in both sample means and variances between system 1 and system 2 (quasi-random) for σ_1. In contrast, no statistical differences were observed in either sample means and variances between system 1 and system 2 (random) for σ_1.

3) Sample Means, μ: For a positional error level of σ_1, the mean of CT for system 1 was detectably higher than that of system 2 (spiral), not different than that of system 2 (random), and detectably higher than that of system 2 (quasi-random). For a positional error level of σ_2, the mean of CT for system 1 was detectably lower than that of system 2 (spiral), not different than that of system 2 (random), and detectably higher than that of system 2 (quasi-random).

4) Sample Variances, σ^2: For a positional error level of σ_1, the variance of CT for system 1 was significantly larger than that of system 2 (quasi-random). By inspection, the variance of CT for system 2 (spiral) appears significantly smaller than that of system 1, but this is statistically not the case. For a positional error level of σ_2, the variance of CT for system 1 is significantly smaller than that of system 2 (spiral), but also significantly larger than that of system 2 (quasi-random).

5) Probability of Success, PS: The probability of inserting a peg for system 1 was statistically equivalent to all variants of system 2 for both levels of positional error despite the fact that system 1 experienced four failures for σ_1 and three failures for σ_2, while system 2 did not experience any failures. Based on KC, system 1 would have to fail at least five times to see a significant difference versus system 2 (assuming system 2 does not fail) at 60 samples. If one wanted to make the claim that the PS values of system 2 are indeed significantly different from those of system 1, then more samples would have to be collected.

TABLE III
PERFORMANCE MEASURES OF ROBOTIC SYSTEMS ON PEG-IN-HOLE TASK WITH POSITIONAL ERRORS σ_1 AND σ_2

Robotic system	Error	KS	μ (s)	σ^2 (s^2)	PS (%)
System 1	σ_1		11.70	40.57	85.4
	σ_2	*	18.31*	107.3*	87.6
System 2 *Spiral*	σ_1		7.19	12.79	95.2
	σ_2	*	37.13*	399.6*	95.2
System 2 *Random*	σ_1		8.01	59.92	95.2
	σ_2	*	15.62*	417.72*	95.2
System 2 *Quasi-Random*	σ_1		3.11	8.84	95.2
	σ_2	*	8.2*	50.25*	95.2

*Statistically significant difference when comparing performance of each system at σ_1 and σ_2.

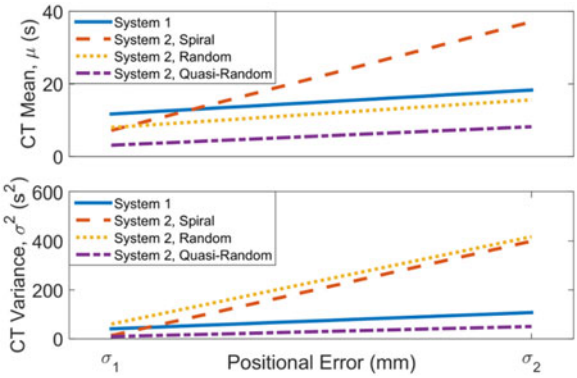

Fig. 8. Performance mean and variance sensitivity to positional error.

B. Within-System Comparison for Positional Sensitivity

Another perspective for assessing peg-in-hole capabilities for a robotic system is considering its performance sensitivity to positional error. In theory, one should be able to execute this peg-in-hole test method at various magnitudes of positional error and create a response curve. The curve would be indicative for how robust a robotic system is at handling positional error for peg-in-hole problems.

The collected datasets are analyzed once more, except statistical comparisons are made within each system at both levels of positional error. The results are reported in Table III. A general remark for these results is that all systems were sensitive to positional error since there were detectable differences for CT sample distributions, means, and variances.

1) Sample Means, μ: All within-system comparisons yielded statistically significant differences in sample means across the two levels of positional error. As shown in Fig. 8, system 2 (spiral) is most sensitive to positional error. The mean of system 2 (spiral) inflates by a factor of 4.6 across positional levels. System 1 is the least sensitive to positional error with its sample mean inflating by a factor of 1.6. System 2 (random) and system 2 (quasi-random) are approximately as sensitive as system 1 as indicated by the near-parallel lines in Fig. 8. At an inflation factor of approximately 2, these systems possess a near 1:1 scaling in performance with positional error.

2) Sample Variances, σ^2: As previously stated, all within-system comparisons yielded statistically significant differences in sample variance across the two levels of positional error. However, Fig. 8 reveals that system 2 (spiral) and system 2 (random) have the greatest variance sensitivity to positional error. In fact, the variance of system 2 (spiral) approximately inflates by a factor of 31 with twice

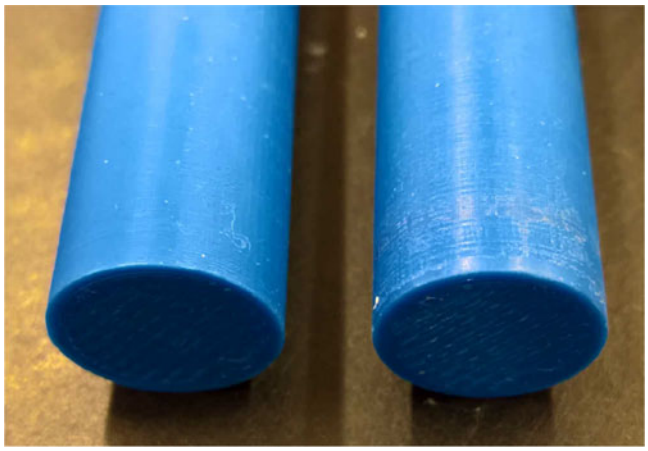

Fig. 9. Surface condition of a peg prior to testing (left) and after testing with system 2 (right).

the level of positional error. On the other hand, system 1 and system 2 (quasi-random) are not nearly as sensitive. The variance of system 1 approximately inflates by a factor of 2.6, indicating a near 1:1 scaling with positional error.

V. Discussion

Assuming that faster CT's are preferential, one can make the following arguments thus far regarding cross-system comparisons at their fixed control parameters for this peg-in-hole task.

1) System 1 and system 2 (random) perform indistinguishably regardless of tested positional error levels;
2) System 1 performs worse than system 2 (spiral) for smaller positional error;
3) System 1 performs better than system 2 (spiral) for larger positional error; and
4) System 1 performs worse than system 2 (quasi-random) regardless of positional error.

When making within-system comparisons, the following results hold. 1) System 2 (spiral) is by far the most sensitive to positional error, and 2) system 1 is the least sensitive to positional error (as it exhibits a near 1:1 scaling with the magnitude of positional variability.)

Although not the focus in this work, other factors worth considering are wear-and-tear and reconfigurability. During testing, more aggressive peg–hole interactions by system 2 were observed than those by system 1. In fact, visible scarring was noticed on the ABS pegs after testing with system 2, as shown in Fig. 9. In contrast, system 1 emitted less than 0.5 N of force between the peg and hole at all times during the insertion operation through force feedback and control. For those applications that require the preservation of surface finishes, robotic systems that perform similarly to system 1 are likely more attractive. Finally, since an element of system 1 is a dexterous, multifingered hand, it can readily be applied to pick-and-place or insertion operations of parts with a greater variety in size and shape without the costs associated with end-effector retooling. This capability is not present with system 2 with a stiff, parallel gripper.

VI. Conclusion

A peg-in-hole test method with associated metrics was created as a performance measurement and benchmarking tool for robotic systems. A thorough statistical analysis procedure is presented to help analyze performance data across and within robotic systems. The first task-level application of a recently developed sensory-rich and complex multifingered control strategy was demonstrated. Applying the peg-in-hole test revealed that the hand system solved this insertion task at a comparable level to a more standard robot system design that leveraged a compliant or force-controlled arm and a simple gripper executing planar search routines. Since parallel grippers are widely used for their simplicity, lower costs, mechanical robustness, and effectiveness for many applications, the pathway to a well-performing robotic hand system begins with performing competitively with a gripper system—a performance point demonstrated in this study.

VII. Future Work

Future efforts will focus on the development of test methods that target a more diversified set of insertion, fastening, and assembly operations using standard parts when available. Positional errors will also be applied across full six-dimensional Cartesian space instead of the planar case investigated in this study. The test methods will be implemented to measure the performance of several force-based robot control systems, including the force-based manipulation controller discussed herein. Performance improvements to robotic hand manipulation are expected as well.

References

[1] N. Correll *et al.*, "Lessons from the amazon picking challenge," arXiv preprint arXiv:1601.05484, 2016.
[2] B. Calli, A. Walsman, A. Singh, S. Srinivasa, P. Abbeel, and A. Dollar, "Benchmarking in manipulation research: Using the Yale-CMU-Berkeley object and model set," *IEEE Robot. Autom. Mag.*, vol. 22, no. 3, pp. 36–52, Sep. 2015.
[3] J. Falco, K. Van Wyk, S. Liu, and S. Carpin, "Grasping the performance: Facilitating replicable performance measures via benchmarking and standardized methodologies," *IEEE Robot. Autom. Mag.*, vol. 22, no. 4, pp. 125–136, Dec. 2015.
[4] J. L. Nevins and D. E. Whitney, *Robot Assembly Research and Its Future Applications*. Boston, MA, USA: Springer, 1979, pp. 275–321.
[5] B. Hannaford, D. A. McAffee, and H. Zak, "Performance evaluation of a six-axis generalized force-reflecting teleoperator," *IEEE Trans. Syst., Man, Cybern.*, vol. 21, no. 3, pp. 620–633, May/Jun. 1991.
[6] Y. Yokokohji, Y. Iida, and T. Yoshikawa, ""Toy problem" as the benchmark test for teleoperation systems," in *Proc. Int. Conf. Intell. Robots Syst.*, 2000, pp. 996–1001.
[7] C. Pacchierotti, F. Chinello, M. Malvezzi, and D. Prattichizzo, "Cutaneous haptic feedback to ensure the stability of robotic teleoperation systems," *Int. J. Robot. Res.*, vol. 34, no. 14, pp. 1773–1787, 2015.
[8] R. P. Khurshid, N. T. Fitter, E. A. Fedalei, and K. J. Kuchenbecker, "Effects of grip-force, contact, and acceleration feedback on a teleoperated pick-and-place task," *IEEE Trans. Haptics*, vol. 10, no. 1, pp. 40–53, Jan./Feb. 2017.
[9] J. G. W. Wildenbeest, D. A. Abbink, C. J. M. Heemskerk, F. C. T. van der Helm, and H. Boessenkool, "The impact of haptic feedback quality on the performance of teleoperated assembly tasks," *IEEE Trans. Haptics*, vol. 6, no. 2, pp. 242–252, Apr./Jun. 2013.
[10] S. A. Dalley, D. A. Bennet, and M. Goldfarb, "Functional assessment of the vanderbilt multigrasp myoelectric hand: A continuing case study," in *Proc. Int. Conf. Eng. Med. Biol. Soc.*, 2014, pp. 6195–6198.
[11] N. Pernalete, W. Yu, R. Dubey, and W. Moreno, "Telerobotic haptic system to assist the performance of occupational therapy tests by motion-impaired users," in *Proc. Int. Conf. Robot. Autom.*, 2003, pp. 1247–1252.
[12] V. Varalta, A. Picelli, C. Fonte, G. Montemezzi, E. La Marchina, and N. Smania, "Effects of contralesional robot-assisted hand training in patients with unilateral spatial neglect following stroke: A case series study," *J. NeuroEng. Rehabil.*, vol. 11, 2014, Art. no. 160.
[13] G. Boothroyd, P. Dewhurst, and W. A. Knight, "Product design for manual assembly," in *Product Design for Manufacture and Assembly*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2002, ch. 3, pp. 73–133.
[14] J. Markdahl, Y. Karayiannidis, X. Hu, and D. Kragic, "Distributed cooperative object attitude manipulation," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2012, pp. 2960–2965.

[15] B. Calli and A. M. Dollar, "Vision-based model predictive control for within-hand precision manipulation with underactuated grippers," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2017, pp. 2839–2845.
[16] R. Boughdiri, H. Nasser, H. Bezine, N. K. M'Sirdi, A. M. Alimi, and A. Naamane, "Dynamic modeling and control of multi-fingered robot hand for grasping task," in *Proc. Int. Symp. Robot. Intell. Sens.*, vol. 41, 2012, pp. 923–931.
[17] M. Li, H. Yin, K. Tahara, and A. Billard, "Learning object-level impedance control for robust grasping and dexterous manipulation," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2014, pp. 6784–6791.
[18] K. V. Wyk, "Grasping and manipulation force control for coordinating multi-manipulator robotic systems with proprioceptive feedback," Ph.D. dissertation, Univ. Florida, Gainesville, FL, USA, 2014.
[19] J. Marvel, M. Shneier, R. Bostelman, and J. Falco, "Multi-robot assembly strategies and metrics," *ACM Comput. Surv.*, submitted for publication.
[20] B. Lotter, *Manufacturing Assembly Handbook*. London, U.K.: Butterworths, 1989.
[21] D. E. Whitney, "Chapter 15: Design for assembly and other ilities," in *Mechanical Assemblies: Their Design, Manufacture, and Role in Product Development*. London, U.K.: Oxford Univ. Press, 2004, pp. 379–391.
[22] C. Favi and M. Germani, "From product architecture to assembly sequence: A method to develop conceptual design for assembly based on interface analysis," in *Proc. Enabling Manufacturing Competitiveness and Economic Sustainability: Proceedings of the 4th International Conference on Changeable, Agile, Reconfigurable and Virtual Production*, Montreal, Canada, 2–5 Oct. 2011, A. H. ElMaraghy, Ed. Berlin, Germany: Springer, 2012, pp. 209–214.
[23] D. Gilliam, S. Leigh, A. Rukhin, and W. Strawderman, "Pass-fail testing: Statistical requirements and interpretations," *J. Res. Nat. Inst. Std. Technol.*, vol. 114, no. 3, pp. 195–199, 2009.
[24] P. Hunter, "Margin of error and confidence levels made simple," 2010. [Online]. Available: https://www.isixsigma.com/tools-templates/sampling-data/margin-error-an d-confidence-levels-made-simple/
[25] NIST/Sematech e-handbook of statistical methods, 2015. [Online]. Available: http://www.itl.nist.gov/div898/handbook/eda/section3/eda35a.htm
[26] W. J. Conover, "A Kolmogorov goodness-of-fit test for discontinuous distributions," *J. Amer. Statist. Assoc.*, vol. 67, pp. 591–596, 1972.
[27] A. Bicchi, J. K. Salisbury, and D. L. Brock, "Contact sensing from force measurements," *Int. J. Robot. Res.*, vol. 12, no. 3, pp. 249–262, 1993.
[28] B. Wang, W. Shi, and Z. Miao, "Confidence analysis of standard deviational ellipse and its extension into higher dimensional Euclidean space," *PLoS One*, vol. 10, no. 3, pp. 1–17, 2015.
[29] D. Deiterding and J. Hendrich, "Probability-based robot search paths," in *Advances in Robotics Research*. New York, NY, USA: Springer, 2008, pp. 31–42.
[30] S. Chhatpar and M. Branicky, "Search strategies for peg-in-hole assemblies with position uncertainty," in *Proc. 2001 IEEE/RSJ Int. Conf. Intell. Robots Syst. Expanding Soc. Role Robot. Next Millennium*, vol. 3, 2001, pp. 1465–1470.
[31] M. A. Habib, M. S. Alam, and N. H. Siddique, "Optimizing coverage performance of multiple random path-planning robots," *Paladyn*, vol. 3, no. 1, pp. 11–22, 2012.
[32] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Quasi-(that is, sub-) random sequences," in *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1992, ch. 7, pp. 309–315.