The Study of Dexterous Hand Manipulation: A Synergy-Based Complexity Index

A. Michael West Jr., Federico Tessari, Margaret Wang, and Neville Hogan, *Member, IEEE*

***Abstract*—In this work we tackle the question of how to analyze and objectively quantify the complexity of a manipulation task. The study investigates the kinematic behavior of the hand joints in three different manipulation tasks of growing complexity: reaching-to-grasp, tool use and piano playing. The collected data were processed to extract the kinematic synergies of the hand by means of singular value decomposition. A novel, unbiased metric to determine hand manipulation complexity was based on the cumulative variance accounted for. This Variance-Accounted-For Complexity Index (VAF-CI) reliably distinguished between different manipulation tasks. Moreover, an unsupervised learning method (k-means clustering) was able to use the index to accurately identify the 3 distinct manipulation tasks. These results may be leveraged to improve the control of biomimetic dexterous robots during manipulation tasks.**

***Index Terms*—Kinematic synergy, dexterity, manipulation, complexity, coordination, piano.**

I. Introduction

The study of robotic grippers, hands and exoskeletons is of great academic and industrial interest due to the need for an intuitive, multi-functional and seamless physical-interaction between machines and their external environment [1]. Multiple studies, spanning more than a century, tried to tackle the design of dexterous end-effectors inspired by the superb manipulation abilities of the human hand [2], [3], [4], [5], [6], [7]. However, manipulation proves to be a challenging problem not only from a design perspective, but also from a control standpoint [1], [8].

The "classical" robotic control techniques—based on rigid body dynamics—provide a sound starting point but struggle when faced with the strong non-linearities of contact-rich dynamics [9]. Two main approaches have emerged. On one hand, several studies tried to solve dexterous manipulation by exploiting learning-based methods [10], [11], [12], [13], [14].

Received 1 July 2024; accepted 5 December 2024. Date of publication 17 January 2025; date of current version 27 February 2025. This article was recommended for publication by Associate Editor D. Guarnera and Editor P. Dario upon evaluation of the reviewers' comments. The work of A. Michael West Jr. was supported in part by the Ford Foundation Fellowship, in part by the Takeda Fellowship, and in part by the Accenture Fellowship. The work of Federico Tessari, Margaret Wang, and Neville Hogan was supported by Eric P. and Evelyn E. Newman Fund. *(Corresponding author: A. Michael West Jr.)*

A. Michael West Jr. and Federico Tessari are with the Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: awest36@jh.edu).

Margaret Wang is with the Department of Mechanical Engineering, and the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Neville Hogan is with the Department of Mechanical Engineering, and the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Digital Object Identifier 10.1109/TMRB.2025.3531006

A complementary approach studies how humans exploit their end-effectors—i.e., their hands—to perform object manipulation. The extraordinary manipulation abilities of humans can provide both inspiration for and intuition about how to achieve dexterous robotic manipulation.

One aspect that has facilitated the study of human manipulation is the investigation of hand synergies [15], [16], [17]. A synergy can be defined as a coordinated action of an ensemble of features. Such features can be described at different levels such as muscular [18], [19], [20] (typically measured through EMGs) or kinematic [15], [16], [17], [20], [21] (typically recorded through joint trajectories or limb motions). A leading theory in human neuromotor control supports the hypothesis that humans exploit synergies to control their dimensionally complex neuromusculoskeletal system [22]. In other words, synergies are a lower dimensional operating space that simplifies the control of a larger set of controllable features.

Human hand synergies have been investigated to simplify robot mechanical design and control [23], [24]. Most studies investigated the role of synergies in simple and common everyday tasks, such as grasping and tool handling [15], [16]. However, the human hand is capable of performing a wide array of tasks with varying degrees of complexity (e.g., from grasping cylindrical objects such as a water bottle to performing complex manipulation tasks such as playing the piano).

With this in mind, one large question remains unanswered: "How can we ensure that the tasks analyzed represent the whole spectrum of complex manipulation activities that humans can perform with their hands?". The answer to this question may lead to a more accurate description of how humans exploit synergies to control their hands during dexterous manipulation. Such comprehension could be leveraged to improve both the design of exoskeletons, prostheses, and general robotic end-effectors as well as the execution of existing synergy-based control strategies.

In this study, we present a new metric to quantify the complexity of hand manipulation tasks. This metric can ensure that the study of hand manipulation incorporates tasks that represent the breadth of the hands' dexterity. To validate this metric, we analyzed three different tasks involving various degrees of hand dexterity: (1) a simple human manipulation task, i.e., reaching to grasp a tool, (2) a more advanced occupational task incorporating tool-use, and (3) a widely recognized advanced manipulation exercise: piano playing.

For each of these tasks, the analysis focused on kinematic synergies of the hand and, in particular, quantification of their complexity based on an index extracted from the cumulative variance-accounted-for (VAF) by all synergies. Here, we introduce the VAF-based complexity index (VAF-CI). The efficacy of this index is statistically validated and its robustness is investigated through a sensitivity analysis. Furthermore, an unsupervised machine learning method, based on k-means clustering, is developed as an unbiased verification of the experimental methodology. The results of this study can be used to ensure that future studies of hand manipulation probe a wider range of the hands' capabilities, and related synergistic behavior.

II. Material and Methods

A. Singular Value Decomposition

Given a data matrix $X \in \mathbb{R}^{n \times m}$, where n denotes the number of observations (e.g., the time evolution of a degree of freedom) and m denotes the number of features (e.g., the number of analyzed DoFs) for a given task, synergies can be extracted using the Singular Value Decomposition (SVD) algorithm:

$$X = U \cdot S \cdot V^T \tag{1}$$

where $U \in \mathbb{R}^{n \times n}$ is an orthonormal matrix whose columns denote the temporal evolution of a given synergy, $S \in \mathbb{R}^{n \times m}$ is diagonal matrix of singular values which give an estimate of the VAF in a given synergy, and $V \in \mathbb{R}^{m \times m}$ is an orthonormal matrix whose columns represent a synergy. For a geometrical interpretation of this, see [25].

The VAF enables us to understand how each extracted synergy represents the variability in the original data set. The VAF of a given, i^{th}, synergy can be computed using:

$$VAF(i) = \frac{\sigma_i^2}{\sum \sigma_i^2} \quad for\ i = 1, \ldots, m \tag{2}$$

where σ_i is the singular value on the i^{th} diagonal element of S. Throughout this paper, the VAF is reported as a decimal ranging from 0 to 1. Intuitively, a synergy with a larger VAF will account for a greater portion of the data variability.

B. Variance-Accounted-For Based Complexity Index

In this work we propose a VAF-based complexity index (VAF-CI). Given a certain experimental task, the VAF-CI can be computed using the following equation based on the trapezoidal integral rule:

$$VAF - CI = 1 - \frac{1}{N_{DoF}} \times \sum_{i=1}^{N_{DoF}-1} \frac{1}{2}\Big(cVAF(i) + cVAF(i+1)\Big) \tag{3}$$

where $cVAF(i) = \sum^{i} VAF(i)$ is the cumulative variance-accounted-for at the i-th synergy and N_{DoF} is the total number of degrees-of-freedom. Alternatively, the midpoint rule, rectangular rule, or Simpson's rule could have been used without affecting the proposed metric. Importantly, in the SVD, the total number of synergies coincides with the total number of degrees-of-freedom. Note that the VAF-CI has been formulated as a metric that can range from 0 to 1.

Mathematically, the VAF-CI is the normalized area above the cumulative VAF curve (See Fig. 3). As this area increases, so does the VAF-CI; for this area to increase, the cumulative VAF-curve must decrease. Conceptually, the VAF-based complexity index increases when more synergies are needed to describe the variance in the data. However, from a physical standpoint, this metric has an upper limit of 0.5. If each degree-of-freedom acts independently, then the contribution of each synergy will be weighted equally. In turn, the cumulative variance-accounted-for will approach a straight line whose slope is $\frac{1}{N_{DoF}}$. The normalized area above this curve will approach a maximum at 0.5.

In this work, we tested the hypothesis that as a task increases in complexity, more individual DoFs are needed to perform the task, and this could be uniquely quantified by the VAF-CI. Importantly, here complexity is defined by the necessity of individualized motion: tasks involving joint coordination are deemed less complex, while those demanding joint individuation are seen as more complex. In sum, the proposed VAF-based complexity index offers an intuitive measure of task complexity, eliminating the need to select the number of significant synergies based on arbitrary variance-accounted-for thresholds.

C. Experimental Setup

To test the VAF-CI, we needed to obtain kinematic hand data for a wide range of tasks with widely recognized different levels of complexity. Specifically we considered the three following manipulation tasks: (1) *reach-to-grasp*, as the most basic manipulation task – every able-bodied human is capable of it, (2) *tool-use*, as a more advanced manipulation – most but not all able-bodied humans are capable of it, and (3) *piano playing*, as one of the most advanced manipulation tasks – only a small percentage of specifically-trained able-bodied humans are capable of it. We want to emphasize that piano playing is a highly dynamic task that requires temporal coordination, joint individuation, and the application of varying forces in a goal-oriented, time specific manner.

1) Wire Harnessing: The first two manipulation tasks (reach-to-grasp and tool-use) were conducted within the framework of wire-harnessing. Fig. 1 shows the mock-up electrical cabinet used for the wire-harnessing experiment.

The study enlisted seven (three women and four men) right-handed subjects, ages ranging from 18 to 28 years, all of whom had no documented history of neurological or musculoskeletal issues. All subjects received a comprehensive briefing on the experimental procedures and signed a consent form approved by MIT's Institutional Review Board.

Each subject executed two distinct tasks. The first task, denoted as the '*reach-to-grasp* experiment', entailed reaching out, grasping, and lifting various tools. Specifically, subjects were instructed to reach for and grip items such as scissors, zip ties, a screwdriver, a wire harness with its branches secured, and a wire harness with its branches unsecured, as if they intended to use them. Each subject repeated this action four times for each object. The concatenation of each of these trials, for a given subject, produced one data matrix, X. The

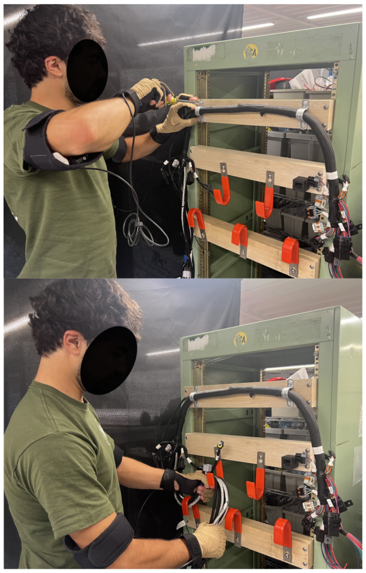

Fig. 1. **Wire harnessing experimental setup:** top image shows the use of a screw-driver (Step 3), bottom image shows the threading of the wire harness (Step 4).

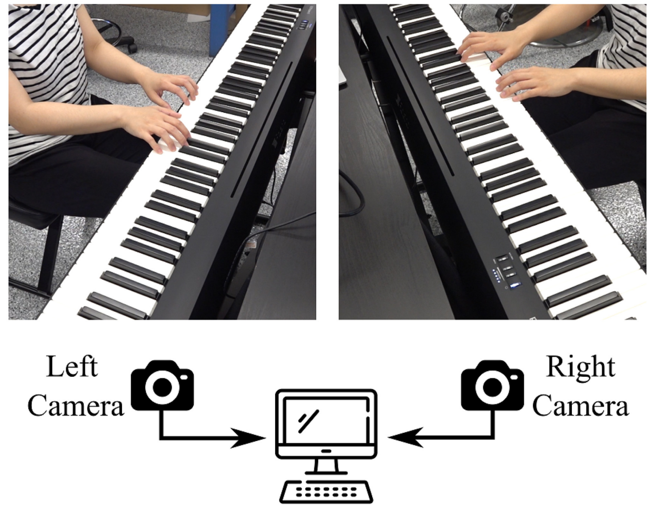

Fig. 2. Piano playing experimental setup.

second task was structured to replicate wire-harness assembly in a manufacturing setting. Subjects were tasked with utilizing the tools from the reach-to-grasp experiment to install a wire harness onto a simulated electrical cabinet (Fig. 1). This task, termed the '*tool-use* experiment', was divided into five distinct steps. At each stage, subjects received verbal instructions and were provided with a booklet containing written directions and a depiction of the completed step. One subject's completion of a given step produced one data matrix X. The steps were as follows: (1) Attach zip ties to the branched ends of the wire harness at three specified locations, (2) employ the scissors to trim excess tips of the zip ties, (3) utilize U-brackets, screws, and a screwdriver to fasten the wire harness securely to the top of the mock electrical cabinet, (4) thread the wire harness through hooks, (5) connect the wire harness into the socket.

Throughout the task performance, we recorded hand postures using a CyberGlove (CyberGlove; Virtual Technologies, Palo Alto, CA, USA), a glove equipped with embedded sensors designed to capture joint kinematics. We recorded 20 joint angles. Specifically, measurements included flexion of the distal interphalangeal (DIP), proximal interphalangeal (PIP), and metacarpophalangeal (MCP) joints of the four fingers. Additionally, abduction (ABD) of the four fingers at the MCP joints was recorded. For the thumb, measurements included flexion at the MCP and interphalangeal (IP) joints, abduction (ABD) at the carpometacarpal joint, and rotation (ROT) around an axis intersecting the trapeziometacarpal joint. During both experiments, subjects wore a CyberGlove on each hand, although only data from the right hand are presented in this report. In all instances, the CyberGlove collected samples at approximately 200 Hz, with a nominal angular resolution of less than 0.1 degrees.

In total, we obtained 10 (5 reach-to-grasp tasks + 5 manipulation tasks) sets of data, $X \in \mathbb{R}^{n\times 20}$ (where $m = 20$ joint DoFs), for each of the 7 subjects. For greater experimental details, see [26].

2) Piano Playing: The piano playing study enlisted five right-handed subjects (three men and two women) with no documented history of neurological or musculoskeletal disorders. Their ages ranged from 19 to 36 years and they had an average piano experience of 20.0 ± 8.3 years. All subjects were right-handed, received a comprehensive briefing on the experimental procedures, and signed a consent form approved by MIT's Institutional Review Board.

Each subject was asked to play any piece they could perform from a series of 9 piano pieces from the classical repertoire, ranging from beginner to advanced level. No subject could play the advanced pieces; thus, they are not reported here: (Beginner) J. S. Bach, *Prelude in C Major, BWV 846*; F. Chopin, *Prelude Op. 28 No. 5*; E. Satie, *Gnossienne No. 1*, (Intermediate) W. A. Mozart, *Sonata No. 8 in A minor, K310, I movement*; L. v. Beethoven, *Sonata No. 23 in F minor "Appassionata", Op. 57, III movement*; C. Debussy, *12 Etudes, No. 5, Pour Les Octaves*.

Afterwards, subjects were invited to play any additional piano piece in their personal repertoire. Each subject was instructed to perform the piece in a unique session without interruptions. This was done to prioritize continuity of motion over accuracy and precision of the piece execution. On average each subject played 6.5 ± 1.9 pieces.

During the execution of every piece subjects hand movements were recorded using two cameras (Sony RX0 II) mounted above the two extremities of the keyboard (Roland FP-10) as shown in Fig. 2. The cameras' recordings were analogically synchronized using two dedicated Sony Camera Control Boxes connected to an Ethernet router. The Sony related acquisition software was used and data were saved at a sampling frequency of 60 *Hz*.

The collected data were then post-processed using the MediaPipe Hands marker-less motion tracking algorithm. MediaPipe Hands is a high prediction quality and lightweight machine learning model. It consists of two modules: (i) a palm detector that places an oriented hand bounding box and (ii) a hand landmark model that operates on the cropped bounding box [27]. This base model detects 21 key-points: one on the wrist and four additional points per finger. A custom algorithm, developed in [28], was used to transform the extracted key-points into 20 joint rotations – the same as those collected by the CyberGlove described above.

The left camera (Fig. 2) was used to extract the right hand middle, ring and little fingers' motion while the right camera took care of the right hand thumb and index fingers. The camera configuration and finger selection were heuristically determined. Only right hand data were analyzed here.

D. Validation of the Variance-Accounted-For Based Complexity Index

Given three sets of tasks with widely accepted different levels of manipulation complexity, we tested the proposed VAF-CI. First the data were centered, in accordance with [25]. Then, after SVD (Eq. (1)), the cumulative VAF (Eq. (2)) and subsequent VAF-CI (Eq. (3)) were computed. Finally, we compared the VAF-CIs of the reach-to-grasp, tool-use, and piano playing trials using a 3 (task) x 1 ANOVA.

1) Comparison to Significant Number of Synergies: To emphasize the robustness of the novel VAF-CI, we compared our metric to the current standard for estimating the number of significant synergies. The literature has used various methods to define the number of significant synergies [15], [17]. Typically, researchers will choose the significant number of synergies using an arbitrarily chosen amount of cumulative VAF. For instance, some have defined the number of significant synergies as the number of synergies required to achieve at least 90% cumulative VAF and where inclusion of another subsequent synergy did not add an additional 5% cumulative VAF [17]. Here, we computed the number of significant synergies for each task considering at least 90%, 80%, or 70% cumulative VAF and where inclusion of another subsequent synergy did not add an additional 5% cumulative VAF. For brevity, each of these three metrics will be referred to as 90 & 5, 80 & 5, and 70 & 5 respectively. For each of the 3 thresholds used to define the significant number of synergies, we performed three one-way ANOVAs (one per each task) to determine if there was an effect of manipulation task on the number of significant synergies.

2) A Metric for Unsupervised Learning: As an additional test, we submitted the computed VAF-CIs to an unsupervised learning algorithm, namely k-means clustering [29], to test whether this metric can be used to classify various manipulation tasks. The k-means clustering algorithm was implemented with the *kmeans* function in MATLAB. Moreover, it was initialized with k centroids whose values were randomly sampled from the input VAF-CI values. To ensure that a local minimum was reached, the algorithm was updated with online updates – each sample was individually reassigned to a new cluster if it reduced the sum of point-to-centroid distances. Finally, in order to find a global minimum this algorithm was repeated 5 times with different randomly-sampled initializations. The repetition with the lowest sum of point-to-centroid distances is reported.

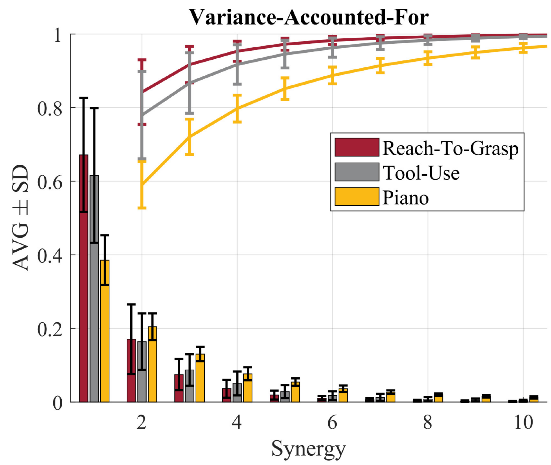

Fig. 3. **VAF of each synergy in the reach-to-grasp, tool-use, and piano playing experiments** averaged across subject. The cumulative sum of the VAF is denoted by the lines. Errorbars are ±1SD. To save space, only the first 10 of 20 synergies are shown.

To identify the number of clusters, we used the elbow method – identifying where adding an additional number of clusters no longer significantly decreases the sum of point-to-centroid-distances. For the case of 3 chosen clusters, we also report the confusion matrix comparing the true and predicted tasks. Moreover, we report the algorithm's accuracy – defined as the number of true positives (correct classifications) for each group divided by the total number of sample tasks.

3) Robustness to Data Pre-Processing: Here, the time needed to complete the various measured tasks could significantly impact the variance in the data, thereby introducing bias to the variance-accounted-for complexity index. To address this concern, we applied two pre-processing techniques before Singular Value Decomposition. Specifically, we downsampled the data using either a specific frequency (60Hz) or a set number of bins (800). These values were chosen to be close to those of the trial with the lowest sampling frequency or number of bins.

As done above, we conducted SVD (Eq. (1)) on the down-sampled data and calculated the cumulative VAF (Eq. (2)) and subsequent VAF-CI (Eq. (3)). Finally, we compared the VAF-CIs of the reach-to-grasp, tool-use, and piano playing trials using a 3 (task) x 1 ANOVA. Also, we report the accuracy of the unsupervised learning algorithm on these data sets when 3 clusters were identified.

III. Results

A. Validation of the Variance-Accounted-For Based Complexity Index

With kinematic hand data from the three sets of tasks (reach-to-grasp, tool-use, and piano playing), we used Eqs. (1) and (2) to compute the variance-accounted-for by each synergy. The average and standard deviation across trials are shown in Fig. 3. Additionally, the cumulative VAFs are shown averaged

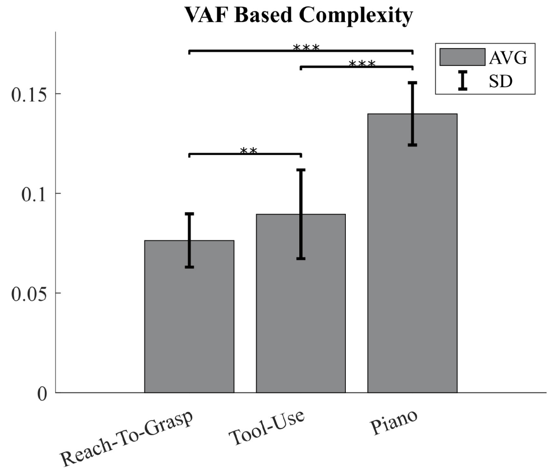

Fig. 4. **VAF-CI in the reach-to-grasp, tool-use, and piano playing experiments** averaged across subject. Errorbars are $\pm$1SD. $*$ represents $p \leq 0.05$, $**$ represents $p \leq 0.01$, and $***$ represents $p \leq 0.001$.

TABLE I
NUMBER OF SIGNIFICANT SYNERGIES FOR THE REACH-TO-GRASP, TOOL-USE, AND PIANO PLAYING EXPERIMENTS TO ACHIEVE AT LEAST 90%, 80%, OR 70% CUMULATIVE VAF AND WHERE INCLUSION OF ANOTHER SUBSEQUENT SYNERGY DID NOT ADD AN ADDITIONAL 5% VAF. THE AVERAGES AND STANDARD DEVIATIONS ARE REPORTED

	90 & 5	80 & 5	70 & 5
Reach-to-Grasp	3.03 ± 1.01	2.83 ± 0.92	2.83 ± 0.92
Tool-Use	3.94 ± 1.35	3.34 ± 0.94	3.34 ± 0.94
Piano	6.97 ± 0.87	4.74 ± 0.71	4.68 ± 0.68

across trials for each set of tasks. The cumulative VAF for a given synergy is generally lower in tool-use than in reach-to-grasp, and lower in piano than in tool-use.

With the cumulative VAF, our proposed VAF-based complexity index was computed using Eq. (3) for each trial. The averages and standard deviations across tasks are reported in Fig. 4. A one-way ANOVA revealed a significant effect of task on VAF-CI ($F_{2,101} = 125.75, p = 3.86e-20$). Post-hoc t-tests (Bonferroni corrected $\alpha = 0.05/3 = 0.0167$) revealed that the tool-use VAF-CI was significantly greater than that of reach-to-grasp ($p = 0.0061$) while the piano playing VAF-CI was significantly greater than that of tool-use ($p = 1.27e-20$) and reach-to-grasp ($p = 2.26e-27$).

B. Comparison to Significant Number of Synergies

Table I reports the number of significant synergies in the reach-to-grasp, tool-use, and piano-playing experiments for the three aforementioned thresholds of cumulative VAF.

Moreover, for each of the 3 metrics used to define the significant number of synergies, we perform a one-way 3 (tasks) x 1 ANOVA to determine if there was an effect of manipulation task on the number of significant synergies. The one-way ANOVA on the 90 & 5 significant synergies returned a significant effect of task ($F_{2,101} = 121.27, p = 1.42e-27$). Post-hoc t-tests (Bonferroni corrected $\alpha = 0.05/3 = 0.0167$) revealed that the tool-use had significantly more synergies than the reach-to-grasp ($p = 0.0021$) while, piano had significantly more synergies than tool-use and reach-to-grasp ($p = 1.48e-19$ and $4.74e-27$ respectively).

In comparison, the ANOVA on the 80 & 5 and 70 & 5 significant synergies demonstrated a significant effect of task in both cases ($F_{2,101} = 44.76, p = 1.20e-14$ and $F_{2,101} = 42.51, p = 4.04e-14$ respectively). However, post-hoc t-tests (Bonferroni corrected $\alpha = 0.05/3 = 0.0167$) revealed that while piano playing exhibited a greater number of significant synergies than tool-use ($p = 3.78e-09$ and $1.13e-08$, respectively) and reach-to-grasp ($p = 1.81e-14$ and $5.35e-14$, respectively) tool-use did not exhibit more significant synergies than reach-to-grasp in both cases ($p = 0.038$ and 0.036, respectively for 80 & 5 and 70 & 5). This demonstrates that arbitrarily choosing a threshold to define which synergies are significant can lead to different results, both in the number of synergies as well as how much they discriminate between manipulation tasks. This method of measuring complexity could lead to arbitrarily biased results.

C. A Metric for Unsupervised Learning

To demonstrate that VAF-CI can be used to classify various manipulation tasks, we submitted these values to an unsupervised learning algorithm, k-means clustering. Fig. 5(a) shows the sum of point-to-centroid distances as a function of the number of clusters. Intuitively, as this value increases to the exact number of samples the sum of point-to-centroid distances will tend to zero. Importantly, Fig. 5(a) shows that after ~3 clusters, adding an additional cluster did not greatly reduce the sum of point-to-centroid distances; there is an "elbow" around 3 clusters. That is consistent with the three different manipulation tasks investigated in this study.

Fig. 5(b) shows the results of the k-means clustering algorithm for three identified clusters. Fig. 5(c) shows the confusion matrix between the identified clusters and the originally considered manipulation tasks. The algorithm correctly identified the task using the VAF-CI with an accuracy of 71.2%. Notably, it classified the reach-to-grasp task with an accuracy of 88.6% and the piano playing task with an accuracy of 82.4%. The tool-use task, whose VAF-CI was generally less than that of the piano-playing task and generally greater than that of the reach-to-grasp task, had a classification accuracy of 42.9%; the tool-use tasks were often classified as a reach-to-grasp task (54.3%).

Fig. 6(a) shows the VAF-CI of the individual tasks that comprise the reach-to-grasp (R-T-G) and tool-use (T-U) experiments. Additionally, the boundaries between clusters identified in Fig. 5(b) are displayed. It is seen that some of the simpler reach-to-grasp tasks, such as plugging in (Step 5) and routing (Step 4) the wire harness, have lower VAF-CIs. On the other hand, the tool use tasks involving more dexterity - using a zip tie and a screwdriver - were classified in the middle cluster.

Fig. 6(b) displays the VAF-CI of the individual piano pieces that were played as part of the piano experiment. Three pieces of generally accepted levels of easy (J.S. Bach, *Prelude in C major, BWV 846* [30]), medium (W.A. Mozart, *Piano Sonata No. 8 in A minor, K. 310*), and advanced (F. Chopin, *Étude Op. 25, No. 11* [31]) difficulty are highlighted. VAF-CI correctly orders these pieces based on complexity, demonstrating that

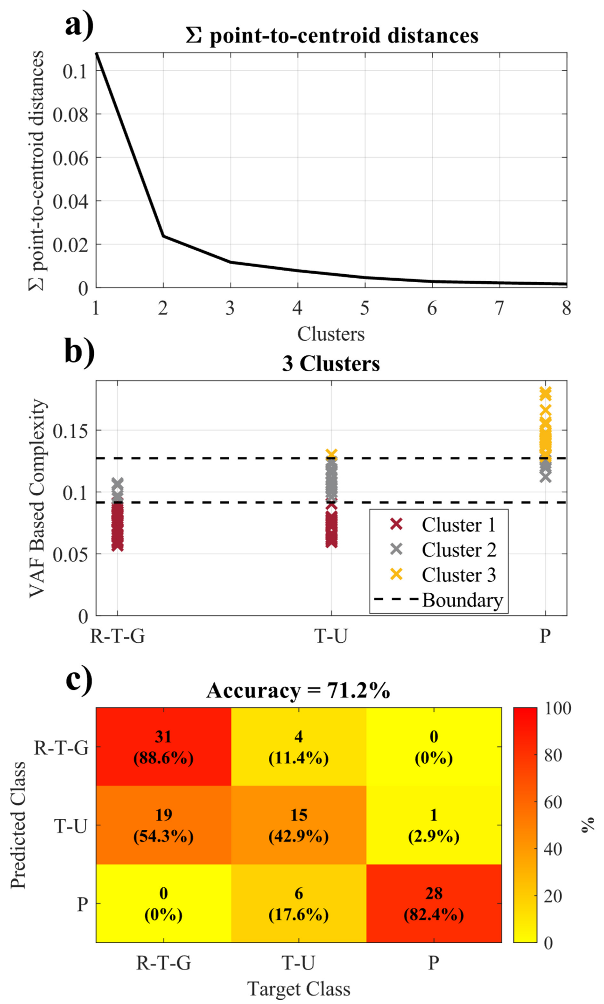

Fig. 5. **K-means clustering algorithm results** (a) The sum of point-to-centroid distances as a function of the number of clusters. At ∼3 clusters, adding another cluster did not greatly reduce the sum of point-to-centroid distances. (b) The identified clusters of the algorithm for 3 chosen clusters and the associated (c) confusion matrix. X's denoted the VAF-CI for each of the samples in the 3 tasks: reach-to-grasp ('R-T-G'), tool-use ('T-U'), and piano playing ('P'). Different colors denote the various identified clusters while the dashed lines denote the boundaries between the identified clusters.

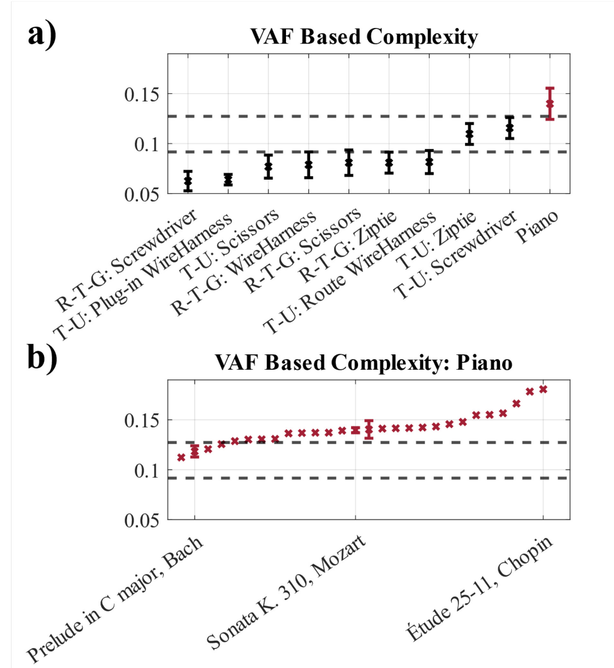

Fig. 6. **Average VAF-CI of individual tasks.** (a) The VAF-CI for the distinct tasks in the reach-to-grasp (R-T-G) experiment and tool-use (T-U) experiment are shown in increasing order in black, while the piano piece VAF-CI is shown in red. (b) The VAF-CI of individual pieces played in the '*piano* experiment' are shown in increasing order in red. Pieces of generally accepted levels of easy (J.S. Bach, *Prelude in C major, BWV 846*), medium (W.A. Mozart, *Piano Sonata No. 8 in A minor, K. 310*), and advanced (F. Chopin, *Étude Op. 25, No. 11*) difficulty are highlighted. Error bars are ±1SD. Note, some piano pieces do not have error bars as they were only played by one subject. Dashed lines denote the boundaries between the identified clusters as in Fig. 5(b).

even within a particular manipulation task (i.e., piano playing), VAF-CI can still capture the varying complexity.

D. Robustness to Data Pre-Processing

As mentioned in Section II-D3, we sought to assess the metric's robustness to data length and pre-processing. Two pre-processing techniques were employed: downsampling to (1) 800 bins and (2) 60 Hz. In the 800 bins data the VAF-CI was 0.076 ± 0.013, 0.089 ± 0.022, and 0.12 ± 0.011 for the reach-to-grasp, tool-use, and piano playing data respectively. Similarly, in the 60 Hz data the VAF-CI was 0.076 ± 0.013, 0.089 ± 0.022, and 0.12 ± 0.012 for the reach-to-grasp, tool-use, and piano playing data respectively. Moreover, in both data sets a one-way ANOVA revealed a significant effect of task on VAF-CI (800 bins: $F_{2,101} = 69.22, p = 1.17e-19$ and 60 Hz: $F_{2,101} = 69.63, p = 9.83e-20$). In both data sets, post-hoc t-tests (Bonferroni corrected $\alpha = 0.05/3 = 0.0167$) revealed that the tool-use VAF-CI was significantly greater than that of reach-to-grasp (800 bins: $p = 0.0030$ and 60 Hz: $p = 0.0035$) while the piano playing VAF-CI was significantly greater than that of tool-use (800 bins: $p = 3.64e-12$ 60 Hz: $p = 2.63e-12$) and reach-to-grasp (800 bins: $p = 1.40e-19$ 60 Hz: $p = 1.25e-19$). Additionally the k-means clustering algorithm, with three identified clusters, had an accuracy of 61.5% and 60.6% for the 800 bins data and 60 Hz data, respectively. Together, these results show that despite varying pre-processing methods altering the data array length, VAF-CI effectively classifies the complexity of the three selected tasks.

IV. Discussion

In this study, we proposed and tested a novel metric for identifying complexity of a manipulation task, termed variance-accounted-for based complexity index (VAF-CI). The VAF-CI is obtained from computing the area above (Eq. (3)) the cumulative VAF-curve (Eq. (2)) obtained after Singular Value Decomposition of the kinematic hand joint data (Eq. (1)). We tested the VAF-CI with respect to three different manipulation tasks of distinct levels of complexity: reach-to-grasp, tool-use, and piano playing. We found a significant effect of manipulation task on VAF-CI (Fig. 3 and 4), validating the index's ability to measure task complexity.

The VAF-CI provides a non-arbitrary metric for quantifying task complexity, differing from previous approaches that relied on the number of synergies to achieve a specific cumulative

VAF [15], [17]. However, this previous approach, is sensitive to the user-defined cumulative VAF cut-off (Table I). While the results at the 90 & 5 threshold align with our VAF-CI metric, it remains uncertain whether this should be considered the standard. Additionally, the enumeration of significant synergies lacks the ability to quantify complexity continuously, as the resulting values are integers. In contrast, VAF-CI eliminates the need for arbitrary choices and enables the quantification of complexity across a spectrum.

Additionally, we demonstrated that VAF-CI, can be used by an unsupervised learning algorithm to identify different manipulation tasks (Fig. 5(b) and 5(c)). K-means clustering was able to identify the 3 different manipulation tasks with an accuracy of 71.2%. Insight can be gleaned from observing how each tool of the wire harnessing task was classified (Fig. 6(a)). The algorithm correctly classified tool-use tasks with an accuracy of only 42.9%. Further inspection revealed that for all 7 subjects, the algorithm classified Step 5, a task involving reaching, grasping and plugging-in an electrical socket, as reach-to-grasp. Given this simple action, this result seems reasonable. Moreover, the algorithm classified scissors tool-use as reach-to-grasp in 6 of the 7 subjects. This may occur because the motion of the joints in the hand is physically constrained by the scissors, leading to motion composed of fewer synergies. Screwdriver reach-to-grasp was always correctly classified, while screwdriver tool-use was classified as tool-use in 6 of the 7 subjects and piano playing in the other. This emphasizes the clear difference in the manipulation complexity required to grasp a screwdriver versus to use it. Notably, this difference was captured by the VAF-CI. Moreover, three of the six piano pieces classified as tool-use were when subjects played J. S. Bach, *Prelude in C Major, BWV 846* – a piece of beginner level [30]. Additionally, Fig. 6(b) demonstrates that even within piano playing there is a spectrum of manipulation complexity and VAF-CI captures this; notoriously complex pieces, such as F. Chopin, *Étude Op. 25, No. 11* [31], were identified as the most complex. Importantly, k-means correctly identified the extremity of piano playing with an accuracy of 82.4% and reach-to-grasp with an accuracy of 88.6%; it never classified piano playing as reach-to-grasp nor reach-to-grasp as piano playing. This emphasizes the robustness of our metric for identifying manipulation tasks without any prior knowledge.

A. Limitations

It is important to emphasize here that manipulation complexity can be composed of many different factors – such as tactile, visual and proprioceptive complexity, and force, torque or stiffness modulation. A limitation of the variance-accounted-for based complexity index is that it only takes into account one factor: joint coordination, or lack-there-of. Nonetheless, it is still a novel and continuous metric that can non-arbitrarily quantify the complexity of different tasks. As such, it is superior to the common practice of arbitrarily identifying an integer number of significant synergies.

While the theoretical limit of the VAF-CI metric approaches 0.5, we only observed a maximum VAF-CI of 0.18 in our study. The theoretical limit is reached when all observed joint motion is completely decoupled. However, this scenario is unlikely in human manipulation tasks, where joint motion is often coupled due to biomechanical constraints. This limitation of biomechanical coupling has been addressed in previous research, such as the findings presented in [17]. Further analysis of their data could help determine the practical upper limit of the VAF-CI metric.

B. Implications

A major challenge in manipulation of dexterous robotic hands is the "curse of dimensionality" [32]. Reducing the control space to a subset of synergies is one solution to this problem [22]. This has led to the design of many synergy-based dexterous robot hands [5], [7], [33], [34], [35], [36]. However, these designs often ignore the fact that synergies may stem from biomechanics as opposed to ease of neuromotor control. In fact, Todorov and Ghahramani [17] found that—when asked to individually actuate each joint in the hand—subjects required significantly more synergies ($\sim$8.7), representing an upper limit to finger individuation and thus manipulation complexity. Contrarily, grasping studies generally report 2-3 significant synergies [15], [16], [20], [21], representing the lower end of finger individuation and thus manipulation complexity. The study presented here introduced the synergy-based analysis of piano playing, a complex manipulation task that requires years of training. Future study of piano playing could provide an upper limit of manipulation complexity during a functional task. Moreover, the distinct difference between reach-to-grasp, tool-use, and piano playing demonstrates the need to study a wide array of manipulation tasks if we truly want to replicate human dexterity. Future work should investigate the synergies that emerge from complex manipulation tasks as building blocks for manipulation.

Perhaps the greatest value of the proposed VAF-CI is as a metric for bio-inspired learning from demonstration. One proposed solution to multi-finger dexterous manipulation is to learn control from human demonstration [37]. However, as previously motivated, if we want our robots to perform dexterous manipulation tasks, we should train them using a rich set of manipulation data. As VAF-CI provides a reliable, continuous, and robust measure of manipulation complexity, it can ensure that a diverse range of manipulation tasks is included in the training data for dexterous manipulators learning from human demonstration, ultimately resulting in improved control policies.

V. Conclusion

Here, we developed a novel complexity index, VAF-CI, able to objectively quantify the complexity of a manipulation task. After measuring the kinematic behavior of the hand joints in three different manipulation tasks of increasing complexity (reaching-to-grasp, tool-use and piano playing), the proposed index identified a clear distinction between these groups. Moreover, it could be used by an unsupervised learning algorithm to reasonably identify 3 different classes of manipulation. This new metric could support the improvement of learning algorithms to achieve robot dexterity from human demonstration and human bio-inspiration. Future work should investigate human dexterity during reasonably complex manipulation tasks quantified by the VAF-CI.

ACKNOWLEDGMENT

The authors would like to thank pianist and composer Alessio Nelli for his support in the selection of the piano pieces.

REFERENCES

[1] A. Billard and D. Kragic, “Trends and challenges in robot manipulation,” *Science*, vol. 364, Jun. 2019, Art. no. eaat8414.

[2] G. Schlesinger, “The mechanical structure of the artificial limbs,” in *Ersatzglieder und Arbeitshilfen*. Berlin, Germany: Springer, 1919, pp. 321–661.

[3] S. Jacobsen, E. Iversen, D. Knutti, R. Johnson, and K. Biggers, “Design of the utah/M.I.T. dextrous hand,” in *Proc. IEEE Int. Conf. Robot. Autom.*, 1986, pp. 1520–1532.

[4] A. Bicchi, “Hands for dexterous manipulation and robust grasping: A difficult road toward simplicity,” *IEEE Trans. Robot. Autom.*, vol. 16, no. 6, pp. 652–662, Dec. 2000.

[5] M. G. Catalano, G. Grioli, E. Farnioli, A. Serio, C. Piazza, and A. Bicchi, “Adaptive synergies for the design and control of the pisa/IIT softhand,” *Int. J. Robot. Res.*, vol. 33, pp. 768–782, Apr. 2014.

[6] C. Piazza, G. Grioli, M. Catalano, and A. Bicchi, “A century of robotic hands,” *Annu. Rev. Control, Robot., Auton. Syst.*, vol. 2, pp. 1–32, May 2019.

[7] M. Laffranchi et al., “The Hannes hand prosthesis replicates the key biological properties of the human hand,” *Sci. Robot.*, vol. 5, p. 467, Sep. 2020.

[8] M. T. Mason, “Toward robotic manipulation,” *Annu. Rev. Control, Robot., Auton. Syst.*, vol. 1, pp. 1–28, May 2018.

[9] R. M. Murray, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*. Boca Raton, FL, USA: CRC Press, Jan. 1994, pp. 1–456.

[10] Y. Li, J. Wu, R. Tedrake, J. B. Tenenbaum, and A. Torralba, “Learning particle dynamics for manipulating rigid bodies, deformable objects, and fluids,” in *Proc. 7th Int. Conf. Learn. Represent. (ICLR)*, Oct. 2018, pp. 1–4.

[11] A. Nagabandi, K. Konolige, S. Levine, and V. Kumar, “Deep dynamics models for learning dexterous manipulation,” in *Proc. 3rd Conf. Robot Learn.*, 2020, pp. 1101–1112.

[12] V. Kumar, E. Todorov, and S. Levine, “Optimal control with learned local models: Application to dexterous manipulation,” in *Proc. IEEE Int. Conf. Robot. Autom.*, 2016, pp. 378–383.

[13] O. Kroemer, S. Niekum, and G. Konidaris, “A review of robot learning for manipulation: Challenges, representations, and algorithms,” *J. Mach. Learn. Res.*, vol. 22, pp. 1–82, Jan. 2021.

[14] O. A. M. Andrychowicz et al., “Learning dexterous in-hand manipulation,” *Int. J. Robot. Res.*, vol. 39, pp. 3–20, Jan. 2020.

[15] M. Santello, M. Flanders, and J. F. Soechting, “Postural hand synergies for tool use,” *J. Neurosci.*, vol. 18, pp. 10105–10115, Dec. 1998.

[16] C. R. Mason, J. E. Gomez, and T. J. Ebner, “Hand synergies during reach-to-grasp,” *J. Neurophysiol.*, vol. 86, no. 6, pp. 2896–2910, 2001.

[17] E. Todorov and Z. Ghahramani, “Analysis of the synergies underlying complex hand manipulation,” in *Proc. 26th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2004, pp. 4637–4640.

[18] A. D’avella and M. C. Tresch, “Modularity in the motor system: Decomposition of muscle patterns as combinations of time-varying synergies,” in *Proc. Adv. Neural Inf. Process. Syst.*, 2001, pp. 1–8.

[19] L. H. Ting and J. L. McKay, “Neuromechanics of muscle synergies for posture and movement,” *Current Opinion Neurobiol.*, vol. 17, pp. 622–628, Dec. 2007.

[20] E. J. Weiss and M. Flanders, “Muscular and postural synergies of the human hand,” *J. Neurophysiol.*, vol. 92, pp. 523–535, Jul. 2004.

[21] M. Santello, M. Flanders, and J. F. Soechting, “Patterns of hand motion during grasping and the influence of sensory guidance,” *J. Neurosci., Off. J. Soc. Neurosci.*, vol. 22, no. 4, pp. 1426–1435, Feb. 2002.

[22] M. Santello et al., “Hand synergies: Integration of robotics and neuroscience for understanding the control of biological and artificial hands,” *Phys. Life Rev.*, vol. 17, pp. 1–23, Jul. 2016.

[23] Z. He and M. Ciocarlie, “Discovering synergies for robot manipulation with multi-task reinforcement learning,” in *Proc. IEEE Int. Conf. Robot. Autom.*, 2022, pp. 2714–2721.

[24] M. T. Ciocarlie and P. K. Allen, “Hand posture subspaces for dexterous robotic grasping,” *Int. J. Robot. Res.*, vol. 28, pp. 851–867, Jun. 2009.

[25] A. M. West Jr. and N. Hogan, “Kinematic hand synergies differ between reach-and-grasp and functional object manipulation,” *Journal of Neurophysiology*, vol. 133, no. 2, pp. 547–560, 2025.

[26] A. M. West Jr. “Towards a non-invasive measurement of human motion, force, and impedance during a complex physical-interaction task: Wire-harnessing,” Master’s thesis, Dept. Mech. Eng., Massachusetts Inst. Technol., Cambridge, MA, USA, 2020.

[27] F. Zhang et al., “Mediapipe hands: On-device real-time hand tracking,” 2020, *arXiv:2006.10214*.

[28] M. Wang, “Non-invasive vision-based measurement of hand kinematics and interaction,” Master’s thesis, Dept. Elect. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, USA, 2023.

[29] S. P. Lloyd, “Least squares quantization in PCM,” *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.

[30] J. S. Bach, *Bach for Beginners*, C. Vincent, Ed. London, U.K.: Boosey & Hawkes Music, Feb. 2011.

[31] F. Chopin and W. Palmer, *Chopin*. Van Nuys, CA, USA: Alfred Publ. Co., Jan. 1995.

[32] R. Bellman, “Dynamic programming and stochastic control processes,” *Inf. Control*, vol. 1, no. 3, pp. 228–239, 1958.

[33] C. Y. Brown and H. H. Asada, “Inter-finger coordination and postural synergies in robot hands via mechanical implementation of principal components analysis,” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2007, pp. 2877–2882.

[34] H. A. Varol, S. A. Dalley, T. E. Wiste, and M. Goldfarb, *Biomimicry and the Design of Multigrasp Transradial Prostheses*. Cham, Switzerland: Springer, 2014, pp. 431–451.

[35] G. C. Matrone, C. Cipriani, M. C. Carrozza, and G. Magenes, “Real-time myoelectric control of a multi-fingered hand prosthesis using principal components analysis,” *J. Neuroeng. Rehabil.*, vol. 9, no. 1, pp. 1–13, Jun. 2012.

[36] A. Furui et al., “A myoelectric prosthetic hand with muscle synergy–based motion determination and impedance model–based biomimetic control,” *Sci. Robot.*, vol. 4, no. 31, p. 467, 2019.

[37] Y. Qin, H. Su, and X. Wang, “From one hand to multiple hands: Imitation learning for dexterous manipulation from single-camera teleoperation,” *IEEE Robot. Autom. Lett.*, vol. 7, no. 4, pp. 10873–10881, Oct. 2022.