In-Situ Measurement of Multi-Axis Torques Applied by Wearable Soft Robots for Shoulder Assistance

Connor M. McCann, *Graduate Student Member, IEEE*, Cameron J. Hohimer, Ciarán T. O'Neill, Harrison T. Young, Katia Bertoldi, and Conor J. Walsh

***Abstract*—While a number of wearable soft robotic devices have been proposed to assist the shoulder, limited efforts have been made to quantify the amount of torque they apply to the body. Most work to-date has assessed soft actuator performance with simple benchtop experiments that may not be representative of the boundary conditions on the human body. We propose a new methodology to measure torques directly in-situ on the body and then use this technique to make a detailed comparison of two versions of a soft wearable shoulder assistance robot. The impact of a number of factors are considered, such as actuator design, garment anchoring, material hysteresis, arm pose, and inflation pressure. Many of these factors are not present on the benchtop and are found to significantly affect torque production. We compare results obtained on a simple benchtop test fixture with two on-body settings: an idealized mannequin and actual human subjects. The mannequin and human results were similar, but differed significantly from the benchtop, further motivating the need for on-body testing. Moving forward, we believe that the ability to directly quantify device performance in-situ will be critical to develop new design, modeling, and control strategies for wearable robots.**

***Index Terms*—Wearable robot, soft pneumatic actuator, textile soft robot, shoulder assistive device, soft robot characterization.**

I. Introduction

OVER the past several years, a significant number of robotic devices have been proposed to assist the shoulder by mechanically restoring or offsetting the natural biological torque of the user's muscles [1]. These devices typically target either medical use cases, seeking to assist clinical populations in a rehabilitation setting [2], [3], or industrial use cases, seeking to aid healthy workers with strenuous overhead tasks [4]. Many proposed systems are constructed with rigid links and joints like a traditional robot, employing electromechanical motors or passive springs to exert forces on the body [5]. One of the key challenges that arises with rigid systems is joint misalignment between the robot and the user, since the gleno-humeral (GH) joint—i.e., the primary ball-and-socket joint where assistance is typically applied—is not stationary, but rather translates with the motion of the scapula [6]. To avoid this difficult misalignment problem, a number of soft robotic devices have recently been proposed that utilize compliance to passively adapt to the complex motion of the GH joint [7]. These soft systems have typically been actuated by Bowden cables that exert tensile forces across the joint [8], [9] or inflatable textile balloons that apply distributed contact forces to the arm [10], [11].

Currently, most systems have been evaluated with biomechanical metrics such as electromyography (EMG) or heart rate measurements that assess the impact of the system on the user [9], [12]. While these techniques do provide relative measures of assistance, they only quantify the response of the human wearer to the device, rather than the actual loads that were applied to elicit that response. To fully understand this causal relationship, direct measurements of the applied loads are required. While most devices apply complex, unknown contact force distributions to the surface of the arm (e.g., through straps, rigid cuffs, or inflatable soft actuators), these loads are typically described as resultant torques about the GH joint for ease of interpretability. Quantitative torque measurements would not only aid in the interpretation of biomechanical data, but would also serve as a valuable benchmarking tool for researchers to compare the assistance provided by different devices. Such measurements would also provide critical ground-truth data for future advances in on-body torque sensing/estimation strategies and advanced numerical soft actuator simulations that consider on-body effects.

Previous efforts have been made to directly measure the actuation loads applied by rigid electromechanical actuators or Bowden cables with integrated torque/force load cells, respectively [13], [14]. To convert these readings to equivalent GH torque values, however, assumptions must be made about the

Manuscript received 23 September 2022; revised 1 February 2023; accepted 2 February 2023. Date of publication 17 March 2023; date of current version 18 May 2023. This article was recommended for publication by Associate Editor M. Munih and Editor P. Dario upon evaluation of the reviewers' comments. This work was supported in part by the National Science Foundation through MRSEC under Grant DMR2011754, through EFRI under Grant 1830896, and through Graduate Research Fellowship under Grant DGE1745303; in part by the Office of Naval Research (ONR) under Award N00014-17-1-2121; in part by the Tata Group; and in part by the Harvard John A. Paulson School of Engineering and Applied Sciences. *(Corresponding author: Conor J. Walsh.)*

This work involved human subjects or animals in its research. The authors confirm that all human/animal subject research procedures and protocols are exempt from review board approval.

Connor M. McCann, Cameron J. Hohimer, Ciarán T. O'Neill, and Harrison T. Young are with the John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138 USA (e-mail: cmccann@g.harvard.edu; cameron.hohimer@gmail.com; ciarnoneill@gmail.com; harrisonyoung@g.harvard.edu).

Katia Bertoldi and Conor J. Walsh are with the John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138 USA, and also with the Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115 USA (e-mail: bertoldi@seas.harvard.edu; walsh@seas.harvard.edu).

This article has supplementary downloadable material available at https://doi.org/10.1109/TMRB.2023.3258499, provided by the authors.

Digital Object Identifier 10.1109/TMRB.2023.3258499

alignment of the sensors relative to the GH joint, resulting in potential inaccuracies, with any off-axis loads being neglected. Moreover, this approach cannot be extended to inflatable soft wearable robots such as [15] and [16] given the distributed nature of their interaction forces. Benchtop test fixtures have previously been utilized to characterize such soft actuators [11], [17], but these boundary conditions are not necessarily representative of those found on the body.

In this work, we present a detailed experimental characterization of two different design variations of an inflatable soft wearable robot, directly measuring the torques applied in-situ on the body. To facilitate these measurements, we propose a new experimental technique for quantifying the applied torques without relying on integrated force/torque sensors, allowing for the characterization of inflatable systems. This method draws inspiration from our prior work, where we presented preliminary on-body torque results by mounting a torque load cell in parallel with the GH joint (much like an isokinetic dynamometer) [18]. This apparatus, however, was prone to misalignment errors and could not adapt to the translation of the GH joint, limiting its utility to static arm trials. In our new approach, we present a different strategy that does not rely on any sensor alignment assumptions, allowing us to measure multi-axis spatial torques. The new technique also applies minimal kinematic constraints to the arm, allowing for dynamic motions without inhibiting the GH joint.

Using this new methodology, we produce a detailed and quantitative performance comparison between the two variations of the soft robotic shoulder support device, illustrating how their behavior is affected by factors such as actuator design, garment anchoring, material hysteresis, arm pose, and inflation pressure. In addition to comparing the two devices, we also investigate the role of boundary conditions by comparing results between a benchtop test fixture, an idealized mannequin, and human subjects. We find the mannequin and human results to be similar, with differences that can be rationalized, whereas the benchtop results are markedly different. This further supports the need for on-body testing of wearable devices.

In Section II, we describe the testing apparatus and experimental procedures. A baseline benchtop comparison of the two soft actuators is then made in Section III. A thorough investigation of on-body performance (using an idealized mannequin) is then presented in Section IV. The generalizability of the technique is demonstrated in Section V with a proof-of-concept human study wherein one of the soft robotic devices is tested on three individuals. Finally, a quantitative comparison of actuator hysteresis across the three boundary conditions is presented in Section VI. Concluding remarks follow in Section VII.

II. Methods

A. Soft Robotic Devices

In this work, a particular class of wearable soft robots for assisting the shoulder are studied, as shown in Fig. 1. These devices employ a "Y"-shaped inflatable textile soft actuator mounted underneath the arm to provide assistance to the wearer. The garment itself is constructed from a combination of knit and woven textiles, forming a harness to react the loads exerted by the actuator. Two previously published versions this device are studied. The first (Device A) was designed for medical applications, aiming to aid impaired individuals after a stroke [15]. The second (Device B) is an iterated version of the system from [16], which aims to assist industrial workers with overhead assembly line work.

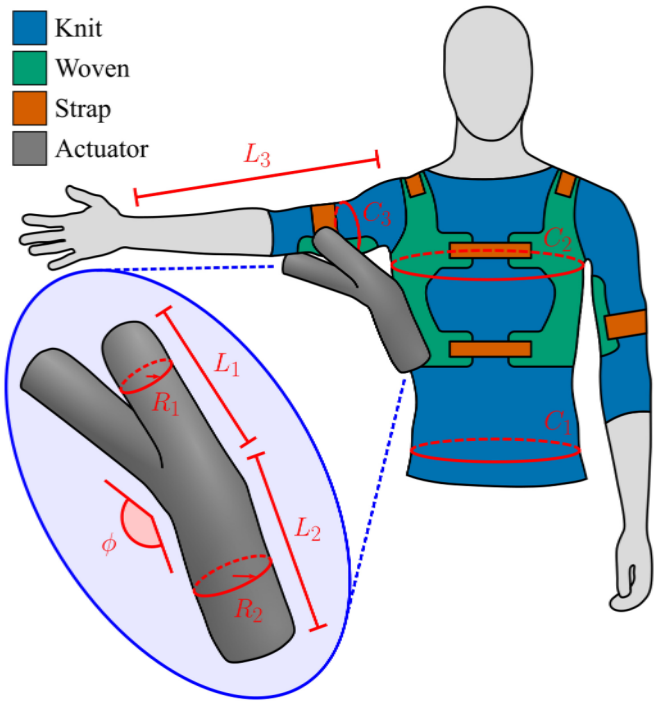

Fig. 1. Schematic of the shirt and soft actuator employed in this work. Two versions of the device have been previously published—for medical and industrial applications—and are examined in this work [15], [16]. The two actuators share the same topology but differ in dimensions, as detailed in Table I. Body dimensions are also labeled, as detailed in Table II.

TABLE I
ACTUATOR DIMENSIONS FOR THE TWO DEVICES CONSIDERED

	Application	L_1	L_2	R_1	R_2 (mm)	ϕ (°)
Device A	Medical	130	165	23	33	165
Device B	Industrial	135	170	19	27	175

TABLE II
BODY DIMENSIONS FOR MANNEQUIN AND HUMAN SUBJECTS

	C_1	C_2	C_3	L_3 (mm)
Mannequin	1005	728	328	486
Participant #1	950	835	270	530
Participant #2	930	735	260	540
Participant #3	845	675	265	500

These two applications inherently gave rise to very different design requirements (see the original papers for further details). For example, for medical tasks, maximizing the amount of assistance was of paramount importance given the impaired nature of the target users, who may not be capable of lifting their arms unassisted. With industrial tasks, however, comfort was of greater importance (since users would wear the device for an entire work day, rather than a shorter clinical therapy session). For this reason—along with the fact that

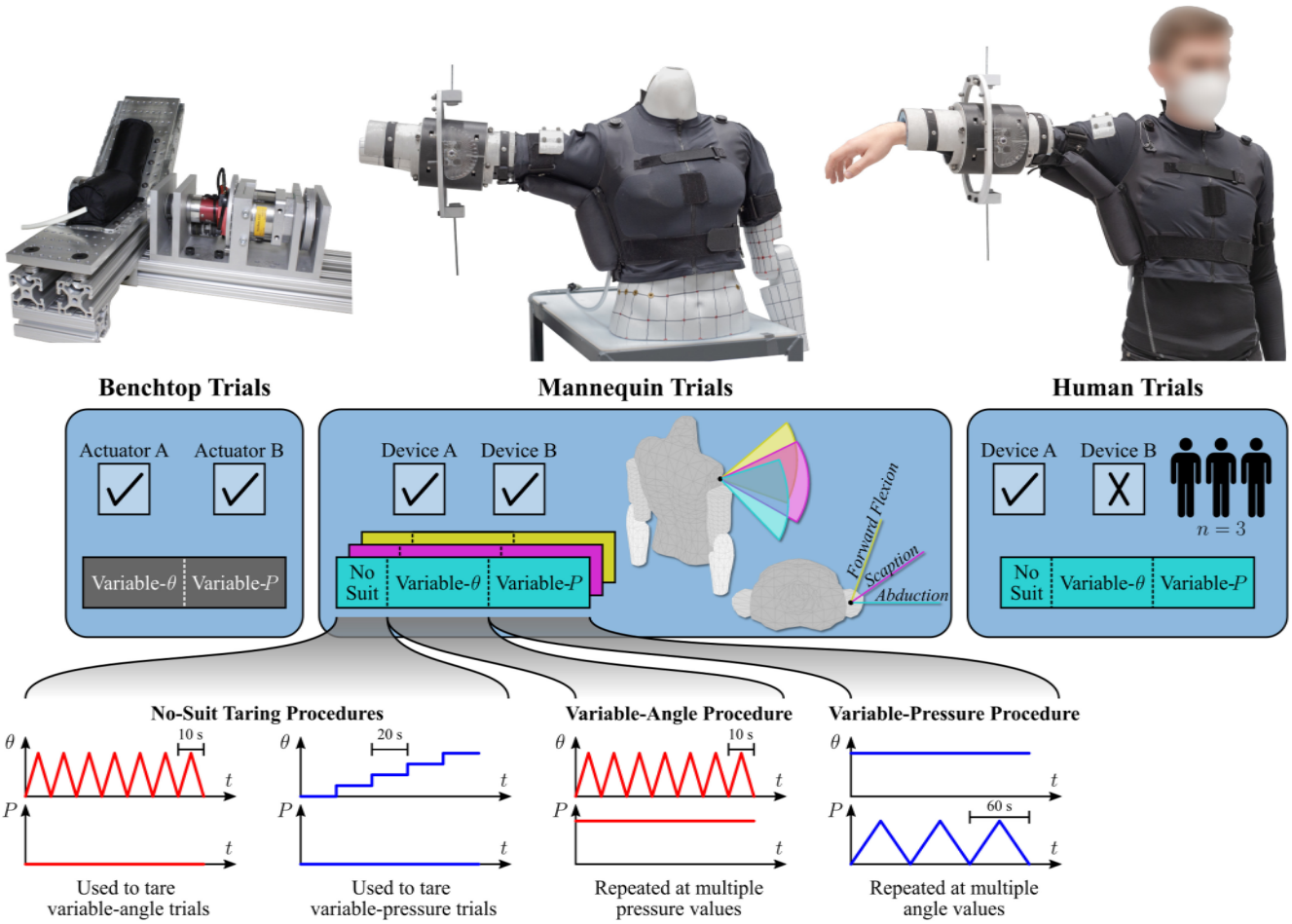

Fig. 2. Overview of experimental test procedures. Both devices were tested on the benchtop and the mannequin (for three horizontal flexion angles, as shown). Proof-of-concept human subject measurements were demonstrated with Device A at one horizontal flexion angle (abduction). For each test condition, both variable-angle and variable-pressure trials were conducted, with the opposite parameter held constant (see inset). For mannequin and human tests, a pair of no-suit trials were also performed to measure the torque due to gravity (and any muscle contribution with the human trials). These no-suit trials were used to tare subsequent measurements.

healthy individuals require less assistance to lift their arms—the torque requirements for Device B were lower. Additionally, assistance in forward flexion was prioritized for Device B, since industrial tasks are most often performed in front of the body, whereas Device A was designed to be more general-purpose. Due to these differing requirements, slightly different actuator and shirt designs were developed for each system. The dimensions of the two actuators are listed in Table I. In this work, we test both systems to study the effect these differences have on their performance, specifically in terms of torque output.

B. Experimental Procedures

To study the effect of boundary conditions, we characterized the performance of these two devices in three settings: a benchtop test fixture, a mannequin, and the human body. These three settings are pictured in Fig. 2, along with an outline of the experimental procedures conducted. The benchtop tests provided a baseline for the performance of the two soft actuators in isolation, but (as will be demonstrated) are not directly representative of the boundary conditions encountered on the body. For the mannequin trials, a poseable mannequin (Zing Displays, USA) was modified with a set of ball bearings to produce a low-friction spherical joint at the shoulder. To ensure an anthropomorphic range of motion, it was necessary to leave a "gap" in the shell of the mannequin surrounding the joint to prevent self-collisions. See Appendix C for experimental validation that this gap does not impact the results obtained. While still "idealized" relative to actual human subjects, this mannequin offers far more realistic boundary conditions than the benchtop fixture. The details of the benchtop and on-body testing setups will be discussed in Sections II-C and II-D, respectively.

To demonstrate the capabilities of the proposed methodology, proof-of-concept tests were performed on three human subjects chosen to have approximately similar body dimensions to the mannequin, for consistency, as shown in Table II. Although achieving an exact dimensional match was not possible (given the highly muscular physique of commercially available mannequins), the qualitative fit of the garment was similar in all cases. While these human tests were the most representative of the final use-case, they also introduced additional complexity and sources of variability such as GH joint translation and muscle stiffness/activation. To completely account for such factors would require a larger participant cohort, and

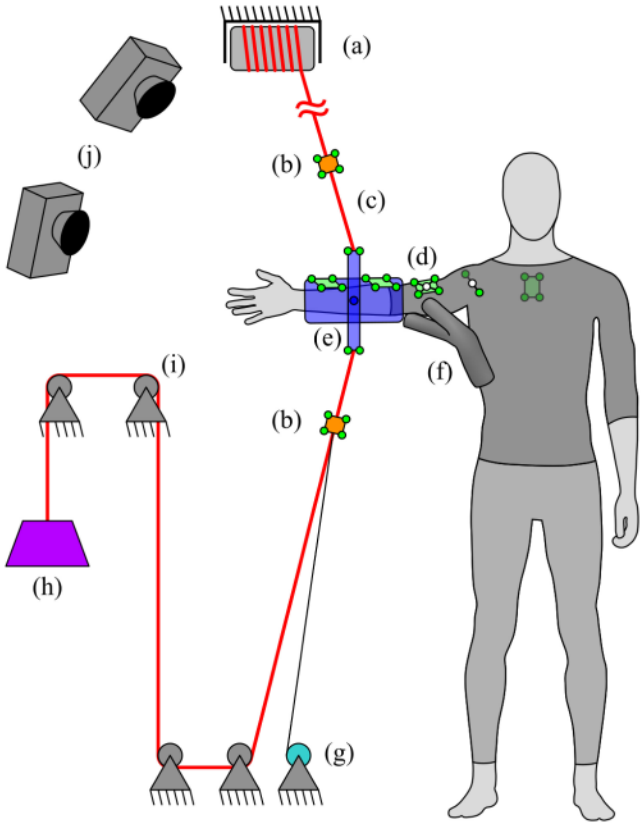

Fig. 3. Schematic of on-body experimental test apparatus used to measure torque in-situ on the body, specifically highlighting: (a) cable winch to adjust arm height; (b) load cells (x2) to "weigh" the arm; (c) load-bearing cable; (d) motion capture markers, with physical/virtual markers shown in green/white, respectively (N.B. the torso marker cluster is on the subject's back); (e) three-axis arm gimbal to isolate the arm from any extraneous torques induced by the cables; (f) soft actuator mounted to a wearable garment; (g) string potentiometer used for real-time control (not utilized in post-processing); (h) preload weight to ensure the cable does not go slack; (i) pulleys to mechanically isolate the arm from any pendulum oscillations of the preload weight; and (j) optical motion capture cameras. Note that the drawing is not to scale.

was beyond the scope of this present work that is focused on describing and demonstrating the potential of the experimental methodology. As such, these human trials were only carried out with Device A. The Institutional Review Board of the Harvard Faculty of Medicine reviewed all human subject testing procedures for this work, and due to the focus of this study on the actuator response rather than any biological effects, this study was determined to be exempt as Not Human Subject Research (NHSR).

In our previous work, the torque produced by these soft actuators was shown to be a function of both pressure and arm angle [18]. For this reason, two testing modes were carried out for each of the three boundary conditions: variable-pressure trials (with arm elevation angle held constant relative to the torso) and variable-angle trials (with internal actuator pressure held constant). In each case, multiple cycles (three for variable-pressure, seven for variable-angle) were performed and then averaged. For the mannequin and human trials, it was necessary to correct for the torque due to gravity with no-suit trials that were used to tare subsequent measurements. With the human trials, muscle activity might also contribute to the torque, and this was also tared using the no-suit trial. This taring procedure are discussed further in Section II-F.

With the mannequin, the horizontal flexion angle of the arm was also varied. Specifically, motions were carried out in front of the torso (forward flexion/extension), diagonally (scapular-plane elevation, or "scaption"), and to the side (abduction/adduction). See Fig. 2 for a visualization of these three directions. Since the human-subject tests were only meant as a proof-of-concept—and because of practical limits on the duration of those trials, as will be discussed in Section V—those trials were limited to side abduction/adduction.

C. Benchtop Testing Apparatus

The actuators were characterized in isolation on a simple benchtop test fixture that was published previously [18]. The fixture consisted of two rigid plates to which the actuators were affixed along their back surface. The plates were joined at a hinge and coupled by a torque load cell to measure the loads exerted by the actuators. For variable-pressure trials, the plates were locked at a fixed angle using an indexing pin. For variable-angle trials, the plates were manually moved between two travel limits with a lever (not pictured in Fig. 2) such that all loads were transmitted through the load cell. For further details on this benchtop testing setup, refer to [18].

D. On-Body Testing Apparatus

The new testing apparatus proposed in this work is shown in Fig. 3. Conceptually, the apparatus works by suspending the arm on a cable from the ceiling. As the soft actuator inflates, it supports a portion of the arm's weight, resulting in a decrease in cable tension. By measuring this tension with a load cell (Futek, USA) and the line of action of the cable with an optical motion capture system (Qualisys, Sweden), a three-dimensional free-body diagram can be constructed, and the reaction torque at the shoulder can be calculated.

To prevent the cable from going slack, a preload weight was suspended below the arm. To avoid undesired pendulum oscillations, the cable supporting this weight was routed through a set of pulleys that prevented horizontal motion. Since these pulleys introduced some friction, a second load cell was used to measure the tension in this bottom cable rather than assuming it to be constant. Given that the two cables may not always be perfectly coaxial, they could produce undesired moments on the arm that would affect the torque balance. To avoid this, a custom-built gimbal was utilized to isolate the arm from any extraneous torques. This gimbal (pictured in Fig. 2) included three intersecting rotational axes and was intentionally designed to immobilize the elbow joint to prevent undesired motions.

In addition to tracking the motion of the gimbal and load cells, the motion capture system was used to track the GH joint location, about which the torque balance is computed. Skin and garment deformation around the shoulder joint make direct marker placement challenging, and thus, the GH joint location was predicted in Qualisys Track Manager (Qualisys, Sweden) as a virtual marker affixed to the upper arm, which was defined as a rigid body via a marker cluster on the arm.

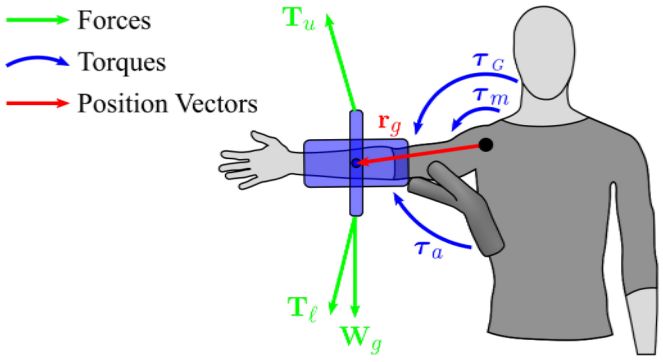

Fig. 4. Free body diagram of the arm, highlighting: the upper and lower cable tensions, $\mathbf{T}_u$ and $\mathbf{T}_\ell$; the weight of the gimbal, $\mathbf{W}_g$; the position vector connecting the glenohumeral joint to the gimbal's center of rotation, $\mathbf{r}_g$; the torque produced by the actuator, $\boldsymbol{\tau}_a$; the torque due to gravity acting on the arm, $\boldsymbol{\tau}_G$; and the torque produced by residual muscle activity in the participant, $\boldsymbol{\tau}_m$, (which can be either positive or negative, depending on whether the subject subconsciously assists or resists the loads being exerted on their arm).

See Appendix A for further details. Additionally, a marker cluster on the participant's back tracked the orientation of the torso. Using the arm and torso orientations, easily interpretable arm angles consistent with typical clinical definitions were calculated using International Society of Biomechanics standards [19].

To allow testing across a range of arm angles, the cable length was adjusted with an electric winch motor on the ceiling. A target computer (Speedgoat, Switzerland) running Simulink Real-Time (Mathworks, USA) was used to control the motor travel based on a string potentiometer signal that served as a proxy for arm height (N.B. the string potentiometer signal was only used for real-time control and did not factor into the data processing). The target machine also logged all sensor signals, controlled actuator inflation, and synchronized measurements with the motion capture system.

E. Actuator Torque Calculations

Using the data obtained from the apparatus, we can construct a torque balance about the GH joint to determine the actuator torque, as depicted in Fig. 4. The force vectors for the upper and lower cable tensions, $\mathbf{T}_u$ and $\mathbf{T}_\ell$, respectively, were determined by multiplying the measured tension values with the unit vectors passing through the load cell centroids and the cable attachment points on the gimbal, as determined from motion capture data. These two forces, along with the weight of the gimbal itself, $\mathbf{W}_g$, which is constant, all acted on the arm at a common point, $\mathbf{r}_g$, defined as the vector from the GH joint to the gimbal center. In addition to these loads, there was also the actuator torque, $\boldsymbol{\tau}_a$, the torque due to gravity, $\boldsymbol{\tau}_G$, and (for humans) any torque due to the subject's muscles, $\boldsymbol{\tau}_m$. The treatment of these last two torques will be discussed further in Section II-F. The actuator torque can be isolated by enforcing static equilibrium between all these loads:

$$\sum \boldsymbol{\tau} = \mathbf{r}_g \times \left(\mathbf{T}_u + \mathbf{T}_\ell + \mathbf{W}_g\right) + \boldsymbol{\tau}_a + \boldsymbol{\tau}_G + \boldsymbol{\tau}_m = \mathbf{0} \quad (1)$$

$$\boldsymbol{\tau}_a = -\mathbf{r}_g \times \left(\mathbf{T}_u + \mathbf{T}_\ell + \mathbf{W}_g\right) - \boldsymbol{\tau}_G - \boldsymbol{\tau}_m. \quad (2)$$

This torque vector provides a full spatial description of the assistance provided. It is also beneficial to calculate a scalar torque quantity for ease of interpretability. In particular, we calculate the component of $\boldsymbol{\tau}_a$ that acts counter to gravity, since this provides an intuitive measure for the amount of assistance provided when raising the arm:

$$\tau = \boldsymbol{\tau}_a \cdot \left(-\hat{\mathbf{n}}_g\right), \quad (3)$$

where $\hat{\mathbf{n}}_g$ is the direction of the torque due to gravity:

$$\hat{\mathbf{n}}_g = \frac{\mathbf{r}_a \times \mathbf{g}}{\|\mathbf{r}_a \times \mathbf{g}\|}. \quad (4)$$

Here $\mathbf{r}_a$ is the vector along the length of the arm and $\mathbf{g}$ is the downward gravity vector. All torque results reported in this work were projected in this way. See Appendix B for a validation of the results obtained with this method relative to our previously published on-body torque testing apparatus.

F. No-Suit Taring Procedure

In order to calculate the actuator torque using Eq. (2), the gravitational and muscle torques, $\boldsymbol{\tau}_G$ and $\boldsymbol{\tau}_m$, must be quantified. The gravitational torque should depend sinusoidally on arm elevation angle, θ (since the arm is essentially a pendulum). The muscle torque, however, is more complicated. All participants were instructed to fully relax their muscles, but it is likely that some muscle activity was still present, as well as some amount of passive muscle stiffness. Moreover, this muscle torque may be positive or negative, depending on whether the subject subconsciously assists or resists the loads being exerted on their arm.

To compensate for these unknown loads, no-suit taring trials were performed before the device was donned to approximate their combined value, $\boldsymbol{\tau}_G + \boldsymbol{\tau}_m$, as a function of arm elevation angle. Since the two appear together in Eq. (2), there is no need to distinguish between them, but the sum can simply be quantified. As shown in Fig. 2, for variable-angle trials, the taring procedure involved sweeping the arm back and forth through a range of elevation angles. This allowed for the characterization of any direction-dependent muscle activation during motion. For variable-pressure trials (when the arm is stationary), this taring procedure consisted of holding the arm at various angles for 20 second intervals and measuring the steady-state torque. In both procedures, the resulting no-suit data were used as an approximation for the combined quantity $\boldsymbol{\tau}_G(\theta) + \boldsymbol{\tau}_m(\theta)$, which was then used in Eq. (2) to tare the readings. All results presented in this work have been tared in this manner.

One inherent limitation of this taring procedure is that the muscle torque, τ_m, is assumed to be only a function of arm elevation angle. In practice, the muscle activity likely fluctuated throughout the experiments. As will be shown in Section V, this led to greater variability in the resulting data. This is one of the main advantages of the idealized testing mannequin, since $\boldsymbol{\tau}_m = 0$, by definition. The human results obtained still appear reasonable despite this limitation.

G. Data Processing

All motion capture data were processed using Qualisys Track Manager (Qualisys, Sweden) and then post-processed in

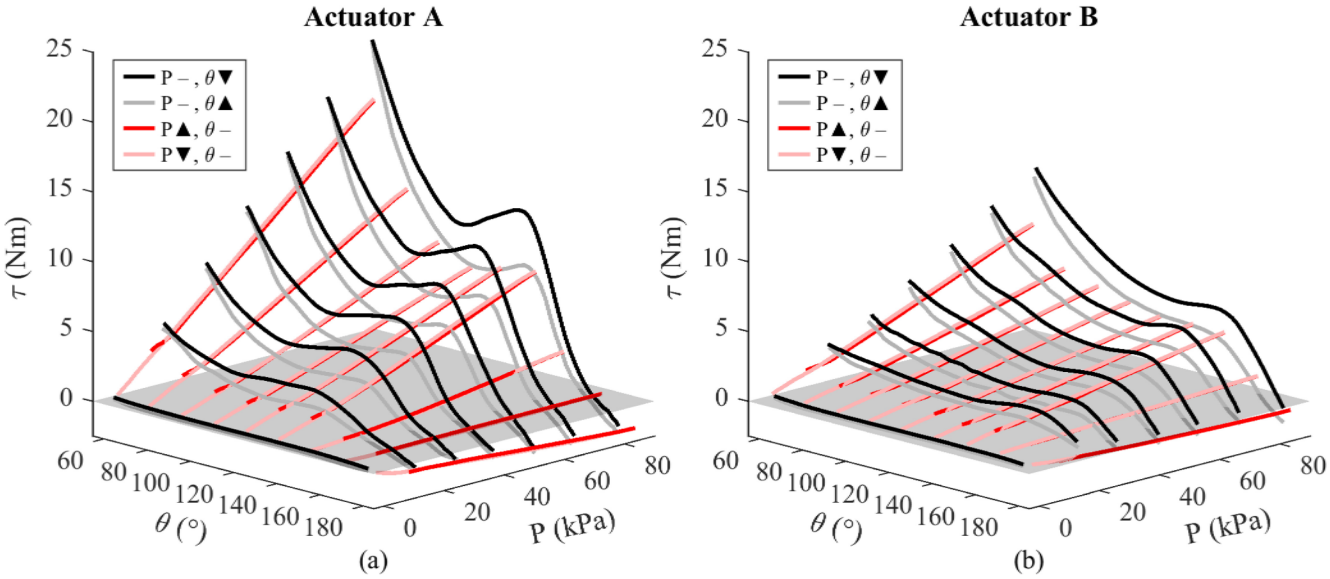

Fig. 5. Soft actuator characterization results using the benchtop test fixture from [18], plotting torque as a function of pressure and the angle between the rigid plates. The actuators from (a) Device A and (b) Device B were tested in isolation, directly anchored to the fixture without the corresponding wearable garments. Hysteresis can be observed in the variable-angle trials (black) but not the variable-pressure trials (red).

MATLAB (Mathworks, USA). The torque load cell data (on the benchtop apparatus) were smoothed with a 5-ms moving average filter. The force load cell data (on the on-body apparatus) were smoothed with an 80-ms fourth-order Savitzky-Golay filter (with the larger window duration to account for mechanical vibrations induced by the winch motor). Data were then segmented into individual cycles and further subdivided into loading and unloading curves based on the appropriate independent variable (i.e., pressure for variable-pressure trials, angle for variable-angle trials). These individual loading/unloading curves were then interpolated to a standardized set of linearly spaced pressure/angle values and the cycles were averaged.

III. Benchtop Results

Benchtop characterization results for the two soft actuators from Devices A and B, respectively, are presented in Fig. 5. These 3D graphs, which will be utilized throughout the paper, depict torque as a function of pressure and angle. Each of the two trial modes—variable-pressure and variable-angle—are plotted, with the darker lines indicating the loading direction (i.e., when external work was done on the system) and the lighter lines indicating unloading.

As can be seen from these results, both actuators responded linearly with pressure but exhibited nonlinear responses as the actuator angle was varied (with Actuator A even exhibiting a non-monotonic torque profile at high pressures in Fig. 5a). As detailed in our prior work, this nonlinearity is due to changes in the self-contacting area of the actuator as it folds over onto itself [18]. Interestingly, hysteresis was observed during variable-angle trials (black) but not during variable-pressure trials (red). We hypothesize that this can be attributed to material hysteresis in the woven textiles, with the individual fibers dissipating energy through friction as they slide over each other. During variable-pressure trials, minimal sliding occurs as the actuators simply expand/contract outward. During variable-angle trials, however, the actuator bends and creases, yielding complex textile deformations that result in sliding of the fibers. The hysteretic response of the actuators will be discussed further in Section VI.

At very high angles, the torque produced by Actuator A became negative. These angles exceeded its pattern angle (165°), and thus the actuator was actually stretched beyond its rest state, resisting motion rather than assisting. Actuator B has a higher pattern angle (175°), and thus did not exhibit significant negative torques. At lower angles ($\theta < 145°$), however, it is evident that Actuator B produced far less torque (∼57% peak magnitude) on the benchtop than Actuator A. This difference is a result of the design decision to prioritize comfort for the industrial use-case (notably the smaller radii, as detailed in Table I). The disparity in torque magnitude was also observed on the mannequin, as will be described in the next section, though the exact nature of this relationship between the two devices becomes more nuanced when on-body factors are considered.

In all cases, the average of multiple cycles (seven for variable-angle trials, three for variable-pressure trials) is plotted in Fig. 5. The results were found to be extremely consistent, with average standard deviations of 0.13 and 0.09 Nm (0.6% and 0.7% peak torque) for Devices A and B, respectively.

IV. Mannequin Results

While the benchtop results presented in the previous section provided a baseline comparison of the two actuators, a more complete understanding requires testing the full wearable robotic systems under more representative boundary conditions. As such, Fig. 6 presents a detailed, side-by-side

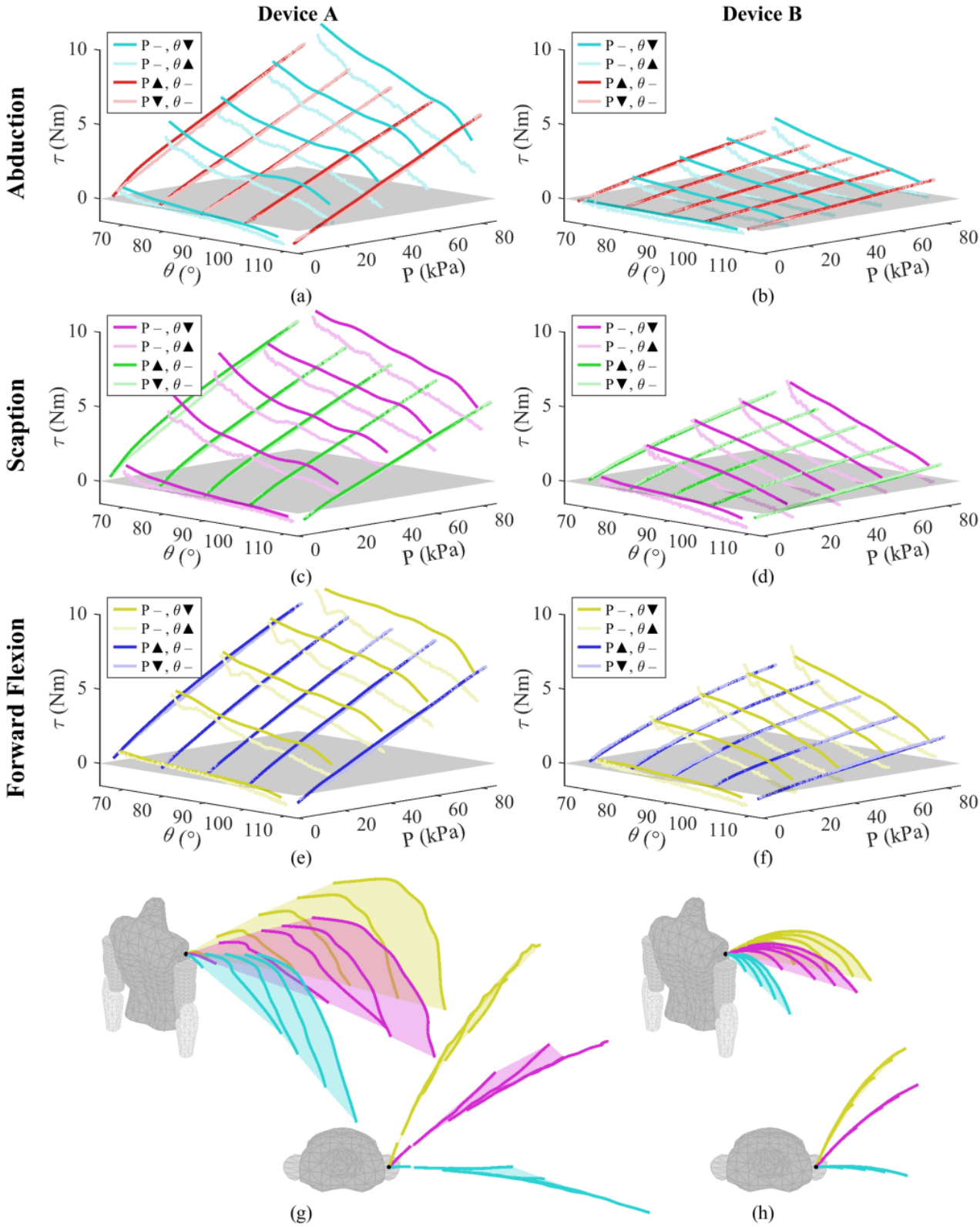

Fig. 6. On-body characterization results for both devices collected on a mannequin. (a)-(f) Torque as a function of arm elevation angle and pressure (for three different horizontal flexion directions). (g)-(h) Polar representation of the same data (with the radius being proportional to torque) plotted relative to the 3D pose of the arm. This provides an intuitive way to visualize the level of assistance available in different parts of the arm's workspace. Two different views are shown for clarity.

comparison of the two devices evaluated on the mannequin. In addition to 3D torque graphs similar to Fig. 5, the results are also visualized in polar form (Figs. 6g and 6h), depicting the amount of assistance across a range of arm poses relative to the mannequin's body, with radius proportional to torque.

A number of comparisons can be made between these mannequin data and the benchtop results. First, a significant drop in torque magnitude was observed, with both devices only exhibiting ∼45% as much peak torque as on the benchtop. This is likely due to the elasticity of the garment providing a less ideal coupling to the body than the rigid constraints of the benchtop fixture. Additionally, the fact that the actuator's center of rotation was not coincident with the GH joint altered the torque magnitude. It is worth noting that Fig. 6 is plotted with respect to arm elevation angle, rather than the benchtop fixture plate angle in Fig. 5. This results in lower angular values, since the actuator angle will generally be greater than the arm angle, as previously reported [18].

The variable-pressure trials still exhibited linear responses with minimal hysteresis. For the variable-angle trials, however, the nonlinearity was mostly eliminated, likely also due to the elastic coupling of the garment that served to smooth out the response (even masking the non-monotonicity observed previously with the Device A on the benchtop). The hysteresis during these trials was of a similar magnitude to that seen on the benchtop, but appears proportionately larger given the reduction in the overall torque values. This larger percent variation with hysteresis may need to be considered when controlling the device, since the amount of assistance provided at a given arm pose may vary significantly depending on the time-history of the user's motions.

As before, the mean of multiple cycles is plotted in Fig. 6. Similar to the benchtop results, limited cycle-to-cycle variability was observed, with average standard deviations of 0.10 and 0.11 Nm (1.0% and 1.1% peak torque) for Devices A and B, respectively. This indicates both that the two devices tested exhibit very low cycle-to-cycle variability for cyclic motions and—more importantly—that the on-body testing apparatus is capable of measuring the torque very repeatedly.

Despite this limited variability within each trial, some trials appear to be outliers relative to the others, such as the 20 kPa variable-angle (teal) trial in Fig. 6a and the 20 kPa and 40 kPa variable-angle (magenta) trials in Fig. 6c, which exhibited elevated torque compared to their neighboring trials at lower angles ($\theta < 80°$). Since the apparatus only constrained the vertical motion of the arm but left the other axes free, there was the potential for the system to reconfigure in unexpected ways. In Fig. 6g, these three trajectories stand out as having rotated slightly toward the front of the body, explaining the discrepancy relative to the other trials.

While the cycle-to-cycle variability for the mannequin data was very low, another factor not captured in these trials is donning and doffing variability. When the device is donned, small variations in the positioning of the actuator can lead it to settle into different configurations when inflated. This type of shifting of the actuator can also occur gradually as the device is worn. To examine the effect of donning and doffing, an additional experiment was performed with Device A in which the 80 kPa variable-angle trial was repeated 12 times, donning and doffing the device between each trial. Care was taken to randomize the order in which the garment's various adjustment straps were tightened to simulate real-world inconsistencies in garment positioning. In this case, the standard deviation of the response increased to 0.84 Nm (8.4% peak torque). The exact amount of donning and doffing variability will be dependent on the specific device being tested. Since the apparatus was able to consistently measure multiple cycles with low cycle-to-cycle variability, this indicates that it could be used to assess the repeatability of different devices, which may be an important performance metric for certain applications where consistent levels of assistance are desirable.

In addition to comparing these mannequin data with the benchtop, comparisons can also be made between the two devices. As seen previously, Device B produced less torque than Device A. In forward flexion, the ratio of peak torques between the two (∼58%) was almost identical to that seen on the benchtop (∼57%) when the two actuators were tested in isolation. When abducting the shoulder, however, response of Device B became attenuated while the performance of Device A remained comparable. This is especially visible in the top views of Figs. 6g and 6h. The difference is likely due to the design decision to prioritize forward assistance with the industrial use-case for Device B by situating the actuator more toward the front of the body, since this is the primary direction in which workers operate, whereas Device A was not designed with this consideration.

Overall, the results in Fig. 6 highlight the need for direct, on-body testing of soft wearable robots, as opposed to relying solely on benchtop testing. Not only did this testing reveal significantly lower torque values on the body, but it also highlighted how different design decisions led to differences in the performance of these two wearable devices (i.e., the dependence on horizontal flexion angle) that were not observable from the benchtop data alone. Moreover, the elasticity of the garment, itself, likely contributed to the decrease in nonlinearities on the body, further motivating the need to assess the entire system in its totality.

V. Human Results

While the mannequin provides an appealing, repeatable testing platform to assess torque production on the body, it is also important to validate that similar results are observed on actual human subjects and to understand any differences that exist between the two. As such, proof-of-concept human trials were carried out (following the same procedure as with the mannequin) with three healthy human subjects (male, ages 23-26) selected to have similar body dimensions to that of the mannequin in order to limit potential compounding factors, as shown in Table II.

The duration of these human subject trials was limited due to the eventual onset of participant discomfort if they were constrained in the testing apparatus for a prolonged period of time. For this reason, these human trials were conducted with only Device A in side abduction/adduction, since these tests could be reliably accomplished before the onset of any discomfort that could potentially bias the data obtained.

The results of these human trials are presented in Fig. 7. Overall, the trends are comparable to those observed on the mannequin (which are reproduced in the figure for a direct

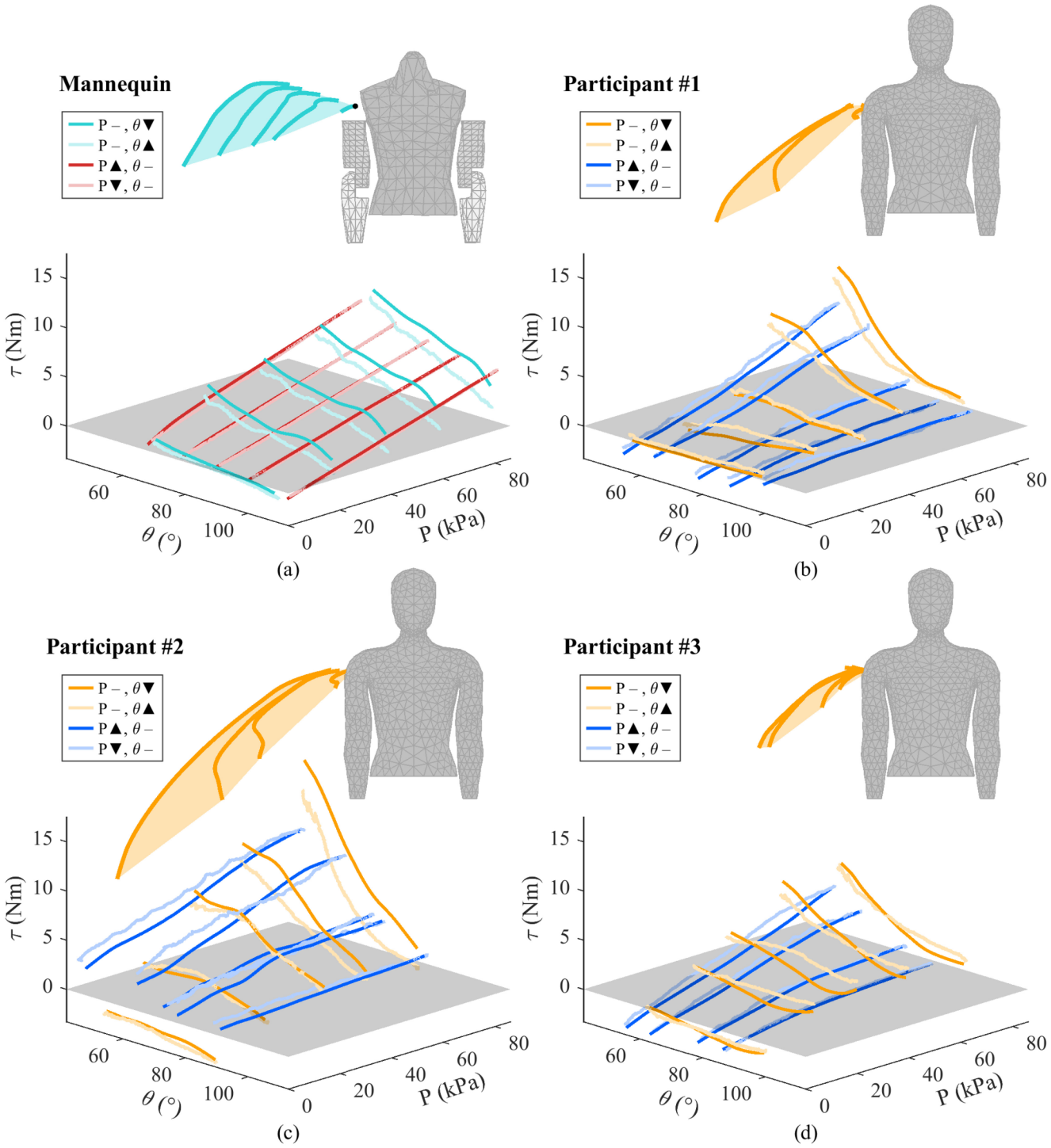

Fig. 7. On-body characterization results for Device A collected on three human subjects in side abduction. Mannequin results are reproduced in (a) for comparison. Despite higher variability and noise in the human data, similar trends and torque magnitudes can be observed as on the mannequin. Attenuation of the torque is observed at higher elevation angles for the human participants, as discussed in Section V.

comparison), albeit with more noise and variability. Moreover, the magnitude of the torque observed was similar to that on the mannequin, unlike the benchtop data, suggesting that the mannequin is a more representative testing platform.

The average cycle-to-cycle standard deviation was 0.61 Nm (5.3% peak torque). Comparing between individuals, the intersubject variability was 1.14 Nm (10.0% peak torque). This was only slightly higher than the 8.4% variability observed with donning/doffing in Section IV, so it is possible that slight differences in the positioning of the garment accounted for the majority of the intersubject variability. It is also possible, however, that personal factors such as body dimensions, soft tissue stiffness, and the degree to which a participant was able to fully relax their muscles also contributed to intersubject variability. Future studies with larger participant cohorts would be required to fully characterize these effects.

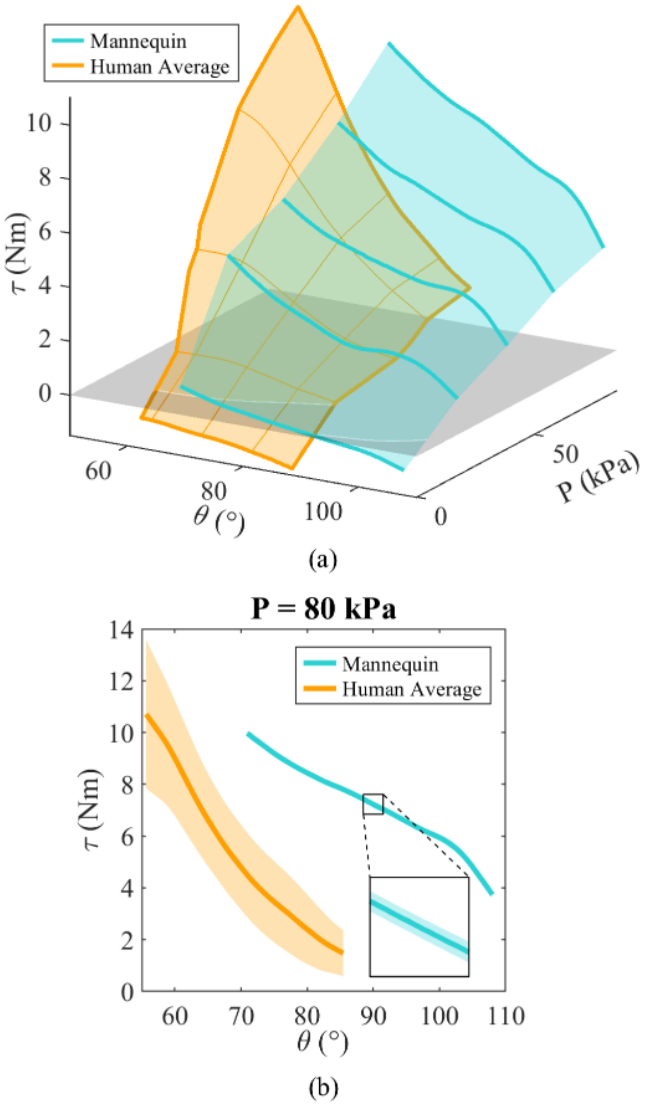

Fig. 8. Comparison of mannequin and human results illustrating the reduction in torque observed at high elevation angles in the human trials. For visual clarity, only the variable-angle loading response is shown, depicted for (a) all measured pressure values and (b) a single pressure value (80 kPa).

One notable difference was an apparent attenuation of the torque produced at higher arm angles compared to the mannequin, as illustrated in Fig. 8 where the average of the three participants' data is plotted against the mannequin results. This trend can likely be explained by the fact that the GH joint translated vertically due to scapular motion with the human participants, whereas it was stationary for the mannequin. This motion, known as the scapulohumeral rhythm, is typically most pronounced at high arm angles [6], and was observed to have an average amplitude of 8 ± 2 cm across the participants. When the GH joint moves upward, it pulls the actuator into the axilla, resulting in the actuator becoming less bent than it would have if no GH joint translation occurred. This, in turn, reduces torque output.

One challenge of human subject testing (and likely the primary source of variability in Fig. 7) is that it is difficult for participants to fully relax their shoulder muscles during these trials. While the taring trials compensated for some level of muscle activation, it is likely that the amount of activity varied trial-to-trial. This further motivates the use of a mannequin to obtain more repeatable results, while human subjects offer greater accuracy to the actual use case. It is likely that a combination of mannequin and human testing is ideal. In future studies that examine a larger cohort of human participants, effects such as garment sizing variability and intersubject variability ought to be investigated further, but these effects were beyond the scope of the present work.

VI. Assessment of Hysteresis

Across all three boundary conditions examined, hysteresis was observed in the variable-angle trials but not in the variable-pressure trials. We hypothesize that this behavior can be explained by frictional energy dissipation within the woven textile of the soft actuator. It has long been understood that the most significant contributing factor to woven-textile hysteresis is yarn-on-yarn friction during textile shearing (when the yarns exhibit the most pronounced sliding relative to each other) [20]. During variable-angle trials, the textile undergoes significant shearing as the actuator creases. During variable-pressure trials, however, the actuator simple expands outward/inward as it inflates/deflates. This biaxial deformation does not result in as much inter-yarn motion, and therefore likely does not cause significant energy dissipation.

To quantitatively compare the hysteresis across the three boundary conditions, we consider the following energy balance for the soft actuator:

$$0 = \tau d\theta + PdV - dD \tag{5}$$

$$D = \int \tau d\theta + \int PdV \tag{6}$$

where τ is the torque, θ is the actuator or arm angle (for the benchtop and on-body tests, respectively), P is the actuator pressure, V is the actuator volume, and D is the energy dissipated due to friction. Generally, to isolate the dissipation term in this expression, the internal volume of the actuator would need to be accurately measured (which can be quite challenging for pneumatic systems due to the compressibility of air). In the case of the variable-angle trials, however, the pressure is held constant, and can therefore be pulled out of the integral. Additionally, since the trials are cyclic, the remaining integral $\oint dV$ is zero, given that there is no net volume change. As such, no net fluidic work is done, and the dissipation over a single cycle is simply given as the area between the loading and unloading torque-angle curves:

$$D = \oint \tau d\theta \tag{7}$$

This hysteric dissipation energy metric was computed for the benchtop, mannequin, and human trials and is plotted as a function of pressure in Fig. 9. For the bencthop and mannequin conditions, the dissipation monotonically increases with pressure (likely because the inter-yarn frictional forces also increase with pressure). The mannequin data exhibited approximately four times less dissipation compared to the benchtop. This is likely because the elasticity of the garment provides a less restrictive set of boundary conditions for the

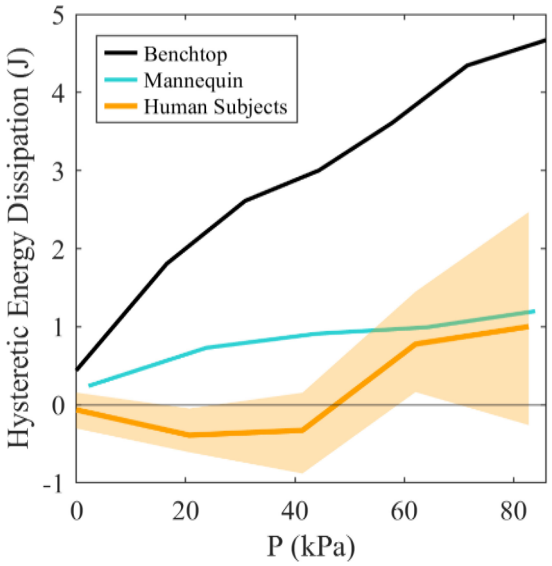

Fig. 9. Comparison of hysteretic energy dissipation during variable-angle trials as a function of pressure, as calculated using Eq. (7). Results are shown for all three testing conditions, with the mannequin and human results calculated for side abduction. The shaded region indicates the minimum/maximum range observed across the three human subjects.

actuator, allowing it to settle into a lower strain energy state compared to the benchtop where it is rigidly affixed.

The human data exhibited significant variability and negative dissipation values. This discrepancy is due to the fact that Eq. (5) does not account for any work done by the human participant's muscles. From these preliminary results, it would seem that the magnitude of the subjects' muscle activity was comparable to the magnitude of the hysteresis, making it difficult to measure hysteresis directly on human subjects. That said, the mean value of the human results appears to approach the mannequin values at higher pressures, which would make sense given that the actuator will undergo very similar deformations between these two on-body cases, and therefore will experience similar amounts of textile shearing. This similarity in dissipation magnitude potentially indicates that he mannequin might be a good testing platform to understand on-body hysteresis more reliably than is possible with human subjects. Future studies with a larger number of participants would be required to confirm this correlation.

VII. Conclusion

In this work, we presented a new technique to characterize the torque produced about the shoulder joint by assistive wearable soft robots. Using this methodology, we quantified and compared the on-body performance of two shoulder systems on an idealized testing mannequin and compared the results with human subjects and a simple benchtop fixture. Similar performance was observed between the mannequin and human trials, but the benchtop results differed significantly, indicating that these boundary conditions were not directly representative of those on the body. Additionally, this new testing approach allowed us to quantitatively demonstrate key performance differences between the two devices based on their respective design decisions—namely that one exhibited a dependence on horizontal flexion angle while the other had a more uniform response. This type of behavior could not be ascertained from benchtop testing, alone. These results support the need for direct, in-situ characterization on the body to ensure higher accuracy to the end use case. We hope that this work will motivate future advances in the design, simulation, and control of wearable soft robots that directly consider on-body factors.

Appendix A
GH Joint Virtual Marker Definition

Given that the skin in the vicinity of the GH joint deforms significantly during arm motion, direct marker placement is challenging when trying to locate the joint center. Instead, a static-hold trial was performed where two temporary markers were situated on either side of the shoulder. The GH joint was assumed to lie at the centroid of these two points. This virtual marker was then affixed to a rigid body for the upper arm using Qualisys Track Manager (Qualisys, Sweden). This rigid body was defined by eight markers on the inner core of the arm gimbal along with a cluster of four markers on the upper arm. Since the deltoid muscles deform during arm motion, this cluster was prone to orientation changes, and thus, only its centroid was used when defining the rigid body. All marker positions are shown in Fig. 3.

Appendix B
Validation of Accuracy

In order to verify the results calculated from the new experimental apparatus, we compared the results with those obtained using the on-body torque measurement setup from [18]. This other apparatus mounts a torque cell in line with the GH joint. As discussed in Section I, this previous apparatus is non-ideal for on-body measurement, being sensitive to misalignment errors and being able to react loads about axes it cannot measure. Due to the lack of other on-body torque measurement setups in the literature, however, it provided a means to cross-validate the results.

Fig. 10 compares results obtained with both testing setups using Device A on the mannequin in side abduction. Special care was taken to precisely align the axis of the torque cell with the mannequin's shoulder joint to limit measurement errors. Since the torque cell apparatus could only measure torques about a single axis, the cable apparatus results were projected onto this same axis (in a manner similar to Eq. (3)). It is worth noting that the boundary conditions were not identical, since the torque cell apparatus fully constrained the arm, while the cable apparatus left it free in horizontal flexion and internal/external rotation. As such, slight discrepancies were observed, especially at higher arm angles. Overall, however, close agreement was still observed between the two—with a root-mean-square error of 0.35 Nm (5% peak torque)—providing confidence in the results.

Appendix C
Impact of Mannequin Shoulder Gap

In order to capture a fully representative range of motion with the idealized testing mannequin, it was necessary to

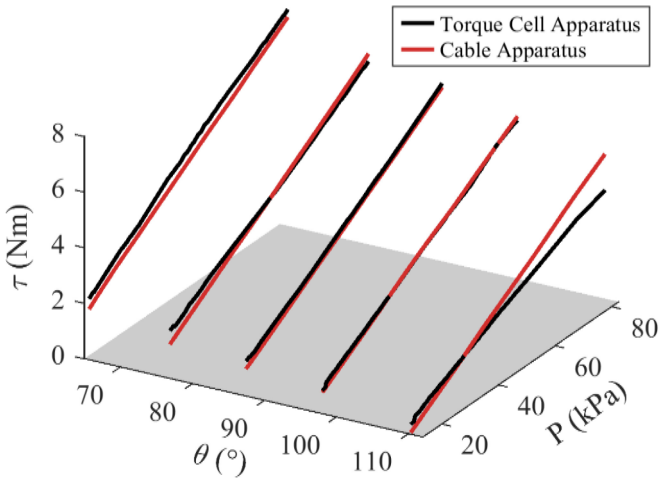

Fig. 10. Comparison between the new cable-based apparatus and the previously published on-body torque cell apparatus from [18]. Results are shown for Device A on the mannequin in side abduction.

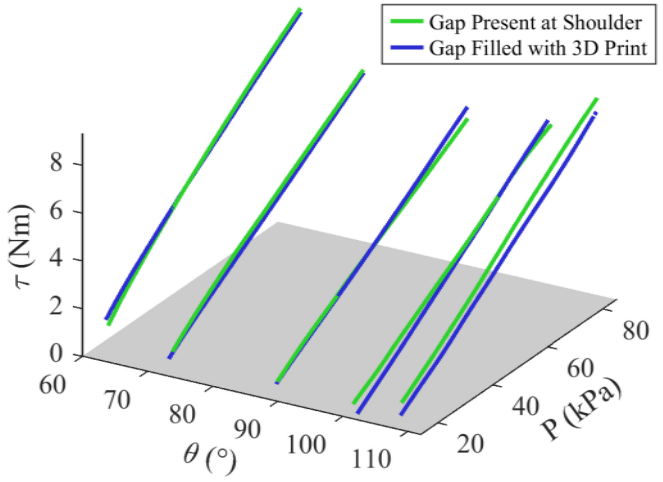

Fig. 11. Comparison between results obtained with the idealized testing mannequin under normal conditions (with a gap at the shoulder joint to prevent self-collisions) and with that gap filled by a 3D printed insert. Results are shown for Device A on the mannequin in side abduction.

remove a portion of the upper arm around the joint—leaving a gap—so as to prevent self-collisions which would invalidate the torque balance assumptions in Fig. 4 and Eq. (2). To ensure that this gap did not bias the results, a set of experiments were performed in which custom-made 3D printed inserts were produced to fill this gap at various static arm elevation angles. A small gap (∼10 mm) was left to prevent self-contact that would invalidate the torque balance assumptions. The results obtained with these inserts is compared in Fig. 11 to those obtained with the gap present. The results were found to be very similar, with a root-mean-square error of only 0.38 Nm (4% peak torque), indicating that the gap has minimal impact on the results.

ACKNOWLEDGMENT

The authors thank Tom Blough and Ted Sirota for their assistance in fabricating components for the setup, Elizabeth Suitor for her help with the electronics, and David Farrell for his insightful discussions throughout this work.

REFERENCES

[1] M. A. Gull, S. Bai, and T. Bak, "A review on design of upper limb exoskeletons," *Robotics*, vol. 9, no. 1, p. 16, 2020.
[2] T. Nef, M. Guidali, and R. Riener, "ARMin III—Arm therapy exoskeleton with an ergonomic shoulder actuation," *Appl. Bionics Biomech.*, vol. 6, no. 2, 2009, Art. no. 962956.
[3] D. Koo, P. H. Chang, M. K. Sohn, and J.-H. Shin, "Shoulder mechanism design of an exoskeleton robot for stroke patient rehabilitation," in *Proc. IEEE Int. Conf. Rehabil. Robot*, 2011, pp. 1–6.
[4] M. P. de Looze, T. Bosch, F. Krause, K. S. Stadler, and L. W. O'Sullivan, "Exoskeletons for industrial application and their potential effects on physical work load," *Ergonomics*, vol. 59, no. 5, pp. 671–681, 2016.
[5] H. M. F. Vatan, S. Nefti-Meziani, S. Davis, Z. Saffari, and H. El-Hussieny, "A review: A comprehensive review of soft and rigid wearable rehabilitation and assistive devices with a focus on the shoulder joint," *J. Intell. Robot. Syst.*, vol. 102, no. 1, p. 9, 2021.
[6] J. Crosbie, S. L. Kilbreath, L. Hollmann, and S. York, "Scapulohumeral rhythm and associated spinal motion," *Clin. Biomech.*, vol. 23, no. 2, pp. 184–192, 2008.
[7] M. Xiloyannis et al., "Soft robotic suits: State of the art, core technologies, and open challenges," *IEEE Trans. Robot.*, vol. 38, no. 3, pp. 1343–1362, Jun. 2022.
[8] I. Galiana, F. L. Hammond, R. D. Howe, and M. B. Popovic, "Wearable soft robotic device for post-stroke shoulder rehabilitation: Identifying misalignments," in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, 2012, pp. 317–322.
[9] S. Lessard, P. Pansodtee, A. Robbins, J. M. Trombadore, S. Kurniawan, and M. Teodorescu, "A soft exosuit for flexible upper-extremity rehabilitation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 8, pp. 1604–1617, Aug. 2018.
[10] C. Simpson, B. Huerta, S. Sketch, M. Lansberg, E. Hawkes, and A. Okamura, "Upper extremity exomuscle for shoulder abduction support," *IEEE Trans. Med. Robot. Bionics*, vol. 2, no. 3, pp. 474–484, Aug. 2020.
[11] C. T. O'Neill, N. S. Phipps, L. Cappello, S. Paganoni, and C. J. Walsh, "A soft wearable robot for the shoulder: Design, characterization, and preliminary testing," in *Proc. IEEE Int. Conf. Rehabil. Robot.*, 2017, pp. 1672–1678.
[12] D. J. Hyun, K. H. Bae, K. J. Kim, S. Nam, and D.-H. Lee, "A lightweight passive upper arm assistive exoskeleton based on multi-linkage spring-energy dissipation mechanism for overhead tasks," *Robot. Autom. Syst.*, vol. 122, Dec. 2019, Art. no. 103309.
[13] D. Park, C. D. Natali, D. G. Caldwell, and J. Ortiz, "Control strategy for shoulder-SideWINDER with kinematics, load estimation, and friction compensation: Preliminary validation," *IEEE Robot. Autom. Lett.*, vol. 7, no. 2, pp. 1278–1283, Apr. 2022.
[14] T. Noda, T. Teramae, B. Ugurlu, and J. Morimoto, "Development of an upper limb exoskeleton powered via pneumatic electric hybrid actuators with bowden cable," in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, 2014, pp. 3573–3578.
[15] T. Proietti et al., "Sensing and control of a multi-joint soft wearable robot for upper-limb assistance and rehabilitation," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 2381–2388, Apr. 2021.
[16] Y. M. Zhou, C. Hohimer, T. Proietti, C. T. O'Neill, and C. J. Walsh, "Kinematics-based control of an inflatable soft wearable robot for assisting the shoulder of industrial workers," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 2155–2162, Apr. 2021.
[17] R. F. Natividad, T. Miller-Jackson, and R. Y. Chen-Hua, "A 2-DOF shoulder exosuit driven by modular, pneumatic, fabric actuators," *IEEE Trans, Med. Robot. Bionics*, vol. 3, no. 1, pp. 166–178, Feb. 2021.
[18] C. T. O'Neill, C. M. McCann, C. J. Hohimer, K. Bertoldi, and C. J. Walsh, "Unfolding textile-based pneumatic actuators for wearable applications," *Soft Robot.*, vol. 9, no. 1, pp. 163–172, 2022.
[19] G. Wu et al., "ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—Part II: Shoulder, elbow, wrist and hand," *J. Biomech.*, vol. 38, no. 5, pp. 981–992, 2005.
[20] W. M. Lo and J. L. Hu, "Shear properties of woven fabrics in various directions," *Textile Res. J.*, vol. 72, no. 5, pp. 383–390, 2002.