Proceedings of Machine Learning Research vol 168:1–16, 2022 4th Annual Conference on Learning for Dynamics and Control

MyoSuite: A contact-rich simulation suite for musculoskeletal motor control

Vittorio Caggiano* CAGGIANO@GMAIL.COM
Meta AI Research, New York, NY, USA

Huawei Wang* H.WANG-2@UTWENTE.NL
and **Guillaume Durandau** G.V.DURANDAU@UTWENTE.NL
and **Massimo Sartori** M.SARTORI@UTWENTE.NL
University of Twente, the Netherlands

Vikash Kumar VIKASHPLUS@GMAIL.COM
Meta AI Research, Pittsburgh, PA, USA

Editors: R. Firoozi, N. Mehr, E. Yel, R. Antonova, J. Bohg, M. Schwager, M. Kochenderfer

Abstract

Embodied agents in continuous control domains have been traditionally exposed to tasks with limited opportunity to explore musculoskeletal details that enable agile and nimble behaviors in biological beings. The sophistication behind bio-musculoskeletal control not only poses new challenges for the learning community but realizing agents embedded in the same perception-action loop that the human sensory-motor system solves can also have a far-reaching impact in fields of neuro-motor disorders, rehabilitation, assistive technologies, as well as collaborative-robotics.

Human biomechanics is a complex multi-joint-multi-actuator musculoskeletal system. The sensory-motor system relies on a range of sensory-contact rich and proprioceptive inputs that define and condition motor actuation required to exhibit intelligent behaviors in the physical world. Current frameworks for studying musculoskeletal control do not include at the same time the needed physiological sophistication of the musculoskeletal systems and support physical world interaction capabilities. In addition, they are neither embedded in complex and skillful motor tasks nor are computationally effective and scalable to study motor learning in the timescale that current learning paradigms require.

To realize a platform where physiological detail and challenges behind human motor control can be investigated, we present a suite of physiologically accurate biomechanical models of elbow, wrist, and hand, with physical contact capabilities which allow complex and skillful contact-rich real-world tasks. The implemented motor tasks provide a great variability of control challenges: from simple postural control to skilled hand-object interactions involving tasks like turning a key, twirling a pen, rotating two balls in one hand, etc. Finally, by supporting physiological alterations in musculoskeletal geometry (tendon transfer), assistive devices (exoskeleton assistance), and muscle contraction dynamics (muscle fatigue, sarcopenia), we present real-life tasks with temporal changes, thereby exposing realistic non-stationary conditions in our tasks which most continuous control benchmarks lack.

Project Webpage: https://sites.google.com/view/myosuite

Keywords: Musculoskeletal model, motor control, neuro-AI

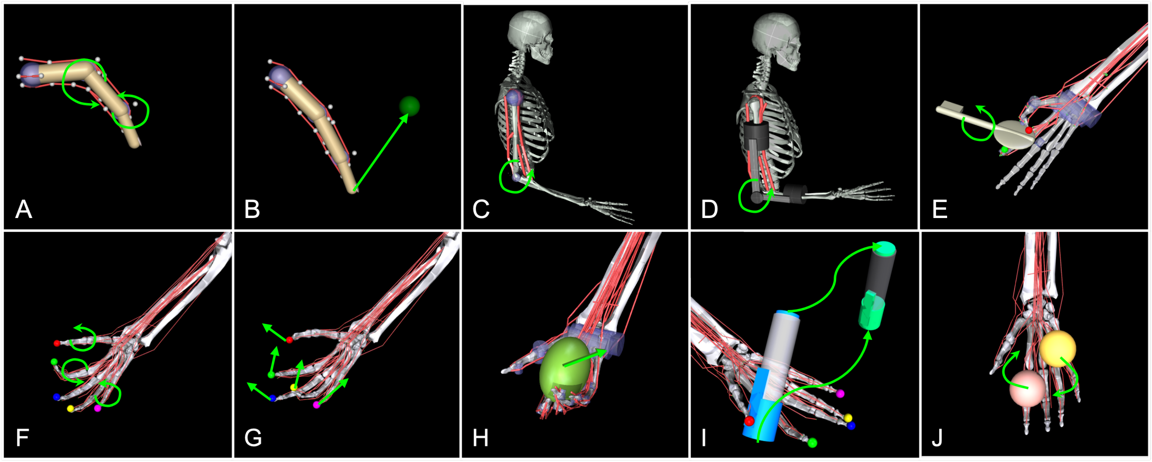

Figure 1: MyoSuite tasks – **Finger**: joint pose (A), tip reach (B). **Elbow**: joint pose (C), exo-assisted (D). **Hand**: key turn (E), joint pose (F), tips reach (G), object reposition (H), pen-twirl (I), baoding balls (J).

1. Introduction

Data-driven learning paradigms have enabled rapid progress in multiple fields such as vision (Deng et al. (2009)), natural language processing (Wang et al. (2018)), speech (Panayotov et al. (2015)), among others. While there has been progress, the field of embodied AI (EAI) is still awaiting its breakthrough moments and real world impacts. Traditionally, challenges in the EAI domain have been proposed by exposing agents to simplified stationary problems that, although have pushed the field forward, do not translate into real-world advancements. One of the root causes is that real-world problems are complex and non-stationary. The human central nervous system, on the other side, can easily handle complex tasks in non-stationary conditions (Latash (2012)) by building control schemes that integrate and model proprioceptive as well as sensory information and transform them into optimal (Shadmehr and Krakauer (2008); Scott (2012)) and adaptable (Wolpert and Ghahramani (2000)) control policies. Such neuro-control challenges that the human central nervous system can seamlessly handle can pose as the next frontier for algorithmic paradigms in the EAI. Furthermore, any development in such a critical problem space can translate into impactful real world advancements in critical fields like neuromechanics, physiotherapy, rehabilitation, as well as robotics.

In comparison to virtual and robotic problems, musculoskeletal control bolsters complexity (task, non-stationarity, as well as dimensionality) that EAI algorithms need to handle. Muscles have third order dynamics and can only generate forces in one direction (pulling). They undergo changes in their force-generating ability depending on their operating length, contractile velocity, or fatigue state (Arnold and Delp (2011)) leading to non-stationarity behaviors. Moreover, muscles may undergo structural changes in response to exercises, aging (e.g. sarcopenia) which further alters their contractile properties over time (Kirkendall and Garrett (1998)). Surgical interventions and assistive devices e.g. exoskeleton, require re-adjustment of their nominal behaviors. In addition, the muscle control space is high dimensional (the number of muscles exceeds the number of human joints - about 600 muscles to control about 300 joints), redundant (multiple muscles act on the same joint), and multi-articular (muscles very often act on multiple joints) (Hirashima and Oya (2016);

Ting et al. (2012); Groote et al. (2016)). As a result, in addition to the non-stationary challenges, the overall system also suffers from the 'curse of dimensionality' (Bernstein (1966)), which is less common in joint level control, which is typical in robotics (Wolpert and Ghahramani (2000)).

The machine learning community has made major advancements by defining and competing on established benchmarks. In the EAI domain, OpenAI-Gym (Brockman et al. (2016)) and DmControl (Tassa et al. (2018)) are the *de-facto* benchmarks for behavior synthesis. These benchmarks however are not suitable for real-world problems. First, they consists mostly of simple problems[1] and are largely already solved. Second, those benchmarks have very limited capabilities to test the adaptability of an agent in response to non-stationary environment changes. Usually, non-stationary settings have been investigated via contrived examples - removing links (Nagabandi et al. (2018)) or adding spurious joints (Gupta et al. (2017); Devin et al. (2017)). There is a dire need for new benchmarks in EAI that expose realistic challenges to the algorithmic paradigms, are embedded closely in the real world, and can translate into real-world impact.

	Contact Dynamics	Muscle Model	Speed
MuJoCo(MyoSuite)	✓	✓	↑
OpenSim		✓	↓
AnyBody		✓	↓
SIMM		✓	↓
PyBullet	✓		↑
Dart	✓	3	↑
RaiSim	✓	3	↑

Table 1: Physic simulators. MyoSuite provides physiologically realistic and computationally efficient musculoskeletal models in the MuJoCo physics engine.

Previous attempts at establishing realistic musculoskeletal tasks as benchmarks (Song et al. (2020)) were narrowly defined. They favored bigger and functionally relevant muscle groups e.g. legs (Hamner et al. (2010); Sartori et al. (2013); White et al. (1989); Ackermann and Van den Bogert (2010)) and arms (Delp et al. (2007); Seth et al. (2018); Saul et al. (2015a); McFarland et al. (2019); Lee et al. (2015); Saul et al. (2015b)). These attempts relied on physics-based musculoskeltal simulation engines such as OpenSim (Seth et al. (2018)), AnyBody (Damsgaard et al. (2006)) and SIM that although are widely used for simulating human neural-mechanical control, human robot interaction, and rehabilitation are computationally expensive (simulating large number of muscle is intractable) and provide limited support for contact rich interactions with their environments(see Table 1). While, physics engines used in the robotic field (PyBullet Coumans and Bai (2016), MuJoCo [2] Todorov et al. (2012), IsaacGym Makoviychuk et al. (2021), RaiSim Hwangbo et al. (2018), and Dart Lee et al. (2018))[3] are relatively more efficient and support contact interactions, but lack adequate support for modeling anatomical and functionally validated musculoskeletal models (see Table 1).

In order to expose the community to the exciting challenges presented by the musculoskeletal control, we present a physiologically realistic and computationally efficient framework: **MyoSuite**.

Our contributions:

- We developed a set of musculoskeletal models (from simple one joint to complex full hand) that are physiologically accurate, several orders of magnitude faster than the state of art musculoskeletal simulators (Ikkala and Hämäläinen (2020); Erez et al. (2015)), and support full contact dynamics.

1. With some exceptions like *HandManipulateEgg* and *HandManipulatePen* (Plappert et al. (2018)) or in-hand manipulation of real-world objects (OpenAI et al. (2019); Huang et al. (2021); Chen et al. (2021))
2. Contains a basic implementation of a muscle model that MyoSuite builds upon
3. Efforts have been made to add muscle models in both DART Lee et al. (2019) and RaiSim Younguk et al. (2018)

- We designed a family of 9 realistic dexterous manipulation tasks (from simple posing to simultaneous manipulation of two Baoding balls) using these models. These task families take inspiration from the state of art robotic dexterous manipulation results (Rajeswaran et al. (2018a); Nagabandi et al. (2020); Andrychowicz et al. (2020)).
- MyoSuite supports physiological alterations in musculoskeletal geometry (tendon transfer, exoskeleton assistance) and muscle contraction dynamics (muscle fatigue, sarcopenia) to expose real-life tasks with temporal changes, thereby subjecting self-adapting algorithms to the realistic non-stationary challenges under continuous control settings.
- In its present form MyoSuite consists of 204 tasks: 9 task-families x 2 difficulty level (easy, hard) x 3 reset conditions (fixed, random, and none[4]) x 8 (or 4) combinations of non-stationarity variations. We present baselines on MyoSuite tasks outlining its features and complexities.

2. Preliminaries

Behavior synthesis for embodied agents can be formulated as a sequential decision-making problem where the goal is to generate a coordinated sequence of trajectories that allows the agent to achieve desired outcomes. Next we outline how behavior synthesis, for musculoskeletal systems in particular, can be formulated using Markov Decision Processes (MDP) formulation under the Reinforcement Learning (RL) paradigm.

2.1. MDP formulation

Per usual notation MDP $\mathcal{M} :: (\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \rho, \gamma)$. $\mathcal{S} \subseteq \mathbb{R}^n$ and $\mathcal{A} \subseteq \mathbb{R}^m$ represent the continuous state and action spaces respectively. The unknown transition dynamics is described by $s' \sim \mathcal{T}(\cdot|s, a)$. $\mathcal{R} : \mathcal{S} \rightarrow [0, R_{\max}]$, $\gamma \in [0, 1)$, and ρ represents the reward function, discount factor, and initial state distribution respectively. Policy is a mapping from states to a probability distribution over actions, i.e. $\pi : \mathcal{S} \rightarrow P(\mathcal{A})$, which is parameterized by θ. The goal of the agent is to learn a policy $\pi^*_\theta(a|s) = argmax_\theta[J(\pi, \mathcal{M})]$, where $J = \max_\theta \mathbb{E}_{s_0 \sim \rho(s), a \sim \pi_\theta(a_t|s_t)}[\sum_t R(s_t, a_t)]$ i.e. the expected sum of discounted rewards in an episodic setting. Policy gradient algorithms (Silver et al. (2014)), TD-learning based algorithms, such as Q-learning (Watkins and Dayan (1992)), SARSA (Sutton (1996)), actor-critic based methods (Konda and Tsitsiklis (2000)), etc. can be leveraged to optimize J to generate behaviors.

2.2. MDP formulation for Robotic systems vs. Musculoskeletal systems

Robotic systems: Under our MDP characterization $\mathcal{M}$, state space $\mathcal{S}$ consists of $\{joint\ position,$ $joint\ velocity\}$, action space $\mathcal{A}$ consists of actuator's $\{position/velocity/torque\ demands\}$ and samples for policy optimization, are gathered either directly from the real world transition $\mathcal{T}_{real}$, or via physics simulation engines $\mathcal{T}_{sim}$ (Todorov et al. (2012); Coumans and Bai (2016); Lee et al. (2018)).

Musculoskeletal systems: Under the MDP characterization $\mathcal{M}$, state space $\mathcal{S}$ consists of $\{muscle-$ $tendon\ length,\ muscle-tendon\ velocity, muscle\ activations\}$, action space $\mathcal{A}$ consists of ac-

4. To facilitate investigation in reset-free algorithms such as Gupta et al. (2021)

tuator's $\{\alpha - Motoneurons\ signals\}$[5] samples for policy optimization, are gathered from the physics simulation engines $\mathcal{T}_{sim}$ controlled via muscle actuators.

Unlike robotic systems, which can be viably investigated by drawing samples from $\mathcal{T}_{sim}$ or $\mathcal{T}_{real}$, owing to the limitations of live experimentation, musculoskeletal simulations $\mathcal{T}_{sim}$ have been the cornerstone of most investigation and understanding behind biological motor control. In contrast to robotics where system identification for comprehensive simulations can be solved comprehensively, mathematical models for biological systems are based on sparse data available from cadavers (e.g. not all parameters can be identified leading to inherent uncertainty).

3. MyoSuite

Unlike robotic counterparts, which are double acting and usually have one-to-one relationships between joints and actuators, tendons and muscles in musculoskeletal systems act via contraction (pull-only) and span multiple joints inducing strong coupling between them. Muscles become fatigued with extended usage. Tendons transfer muscle forces to bones while serving as temporary energy storage units for motion efficiency. These details, while complex, conceal within themselves the ingredients of effective motor control in biological systems. To facilitate investigations in these details, we present the *"MyoSuite"* – a collection of physiologically accurate musculoskeletal models (Section 3.1) and physically realistic contact-rich tasks (Section 3.2) of varying complexity.

3.1. Musculoskeletal models

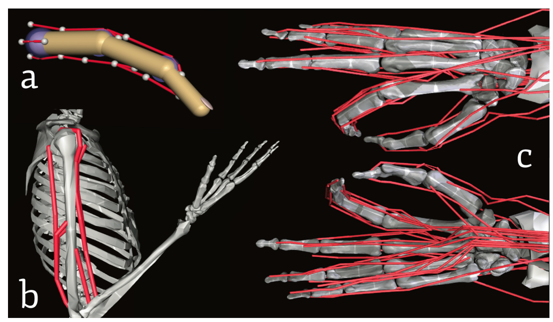

Figure 2: Musculoskeletal models. A - MyoFinger (4 joints - 5 muscles), B - MyoElbow: (1 joint - 6 muscles), C - MyoHand: (23 joints - 38 muscles).

Musculoskeletal models are commonly modeled as a 3rd order system which contains first (or second-order) muscle activation and contractile dynamics (full detail in Online Appendix/Models) as well as the second-order skeletal dynamics. Our models are developed using MuJoCo physics engine Todorov et al. (2012) via a thorough investigation of well studied existing models (Delp et al. (2007); Seth et al. (2018); Saul et al. (2015a); McFarland et al. (2019); Lee et al. (2015); Saul et al. (2015b)) and functional studies (Wu et al. (2008)). We started from OpenSim models (Lee et al. (2015); McFarland et al. (2019); Saul et al. (2015a)) of the arm and hand which are widely used in fields of human neural-mechanical control, human-robot interaction, and rehabilitation. In order to implement those models in MuJoCo, we developed a pipeline Wang et al. (2022) to perform geometry transformations of bones and muscles attachment, moment arm optimization, and muscle force optimization. After rigorous modeling and validation (Section 4.1), we built three physiologically realistic models (Figure 2) of varying complexities.

MyoFinger: A simplified and intuitive model (based on Xu et al. (2012)) of a 4 Degree of Freedom (DoF) finger (MyoFinger, Figure 2A), which is actuated through a series of 5 simplified antagonistic muscle-tendon units. We also provide its robotic counterpart with simple torque actuators to facilitate comparative investigations.

5. Motorneurons are the final neuronal stage that connects the central nervous system to the muscles.

MyoElbow: A model of 1 DoF human elbow joint (Figure 2B). MyoElbow model is based on OpenSim's default testing arm model (Delp et al. (2007); Seth et al. (2018)) and actuated using multiple agonist/antagonist pairs (3 flexors and 3 extensors).

MyoHand: The dexterous human hand is comprised of 29 bones, 23 joints, and 38 muscles-tendon[6] units (see Online Appendix/Models for a detailed description of the hand muscles). This forearm-wrist-hand model (MyoHand, Figure 2C) combined and extended popular OpenSim models: MoBL - human upper extremity model (Saul et al. (2015a) McFarland et al. (2019)) - and the 2nd-Hand - for hand and fingers models (Lee et al. (2015)).

3.2. Tasks

Leveraging these musculoskeletal models, we built a series of tasks (see Figure 1) of varying difficulty. The task difficulty is varied along two axes: task-complexity, and task-non-stationarity. Task-complexity has two variations (*difficulty* - easy/hard, and *Reset* - Fix/Random/None), and task-non-stationarity has 8 (or 4 if tendon-transfer and exoskeleton assistance are not possible) variations. In total, in its current form MyoSuite consists of 204 tasks (Table 2): 9 task-families x 2 difficulties-levels x 3 Resets x 8 (or 4) combinations of non-stationarity variations. We provide a complete description of the tasks and task-complexities in Online Appendix/Tasks. Next we detail various non-stationarities that are supported in MyoSuite tasks.

	Complexity		**Non-Stationarity**				
	Easy/Hard	Reset	None	Sarcopenia	Fatigue	Tendon-transf.	Exo.
Finger Joint Pose	✓/✓	F/R/N	✓	✓	✓		
Finger Tip Reach	✓/✓	F/R/N	✓	✓	✓		
Elbow Joint Pose	-/✓	F/R/N	✓	✓	✓		✓
Hand Key Turn	✓/✓	F/R/N	✓	✓	✓	✓	
Hand Joints Pose	✓/✓	F/R/N	✓	✓	✓	✓	
Hand Tips Reach	✓/✓	F/R/N	✓	✓	✓	✓	
Hand Object Hold	✓/✓	F/R/N	✓	✓	✓	✓	
Hand Pen Twirl	✓/✓	F/R/N	✓	✓	✓	✓	
Hand Baoding Balls	✓/✓	F/R/N	✓	✓	✓	✓	

Table 2: MyoSuite tasks with (a) complexity variations (easy/hard and Reset - **F**ix, **R**andom, **N**one), and (b) non-stationarities variations (None, Sarcopenia, Fatigue, Tendon-transfer and exoskeleton)

3.3. Realistic non-stationary task-variations

Muscle properties are constantly changing. These changes can be instantaneous - like for musculoskeletal injury or surgery - or can vary over a short time frame - like muscle fatigue or exoskeleton assistance. To study neuromuscular adaptation to non-stationarities due to these changes during the real work-life scenario, four different variations in muscle properties have been included: Sacropenia, Fatigue, Tendon Transfer, and Exoskeleton assistance.

Sarcopenia: Sarcopenia is a muscle disorder that occurs commonly in the elderly population (Cruz-Jentoft and Sayer (2019)) and characterized by a reduction in muscle mass or volume. The peak in grip strength can be reduced up to 50% from age 20 to 40 (Dodds et al. (2016)). We modeled sarcopenia for each muscle as a reduction of 50% of its maximal isometric force.

6. We also include an Opponens Pollicis muscle for the critical role it has in manual dexterity (Karakostis et al. (2021))

Fatigue: Muscle Fatigue is a short-term (second to minutes) effect that happens after sustained or repetitive voluntary movement. It has also been linked to traumas e.g. cumulative trauma disorder (Chaffin et al. (2006)). A dynamic muscle fatigue model (Ma et al. (2009)), that builds on the idea that different types of muscle fiber have different contributions and resistances to fatigue (Vøllestad (1997)), was developed and integrated into the modeling framework. See Online Appendix/Non-Stationary for details on the implementation of the fatigue model.

Tendon tear/ tendon transfer via surgery: While muscle fatigue and sarcopenia affects all muscles, accidents can lead to damage of a subset of muscle-tendon units (termed as tendon-tear). Tendon transfer surgery allows redirecting the application point of muscle forces from one location to another (see Figure 3). It can be used to regain functional control of a joint or limb motion after injury. One such tendon transfer procedure is relocation of Extensor Indicis Proprius (EIP) to replace the Extensor Pollicis Longus (EPL) (Gelb (1995)). Rupture of the EPL can happen after a broken wrist. It results in a loss of control over the thumb extension. We introduce a physical tendon transfer where the EIP application point of the tendon was moved from the index to the thumb and the EPL was removed (see Figure 3).

Exoskeleton assistance: Exoskeleton assisted rehabilitation is becoming common (Jezernik et al. (2003)) and demonstrating significant benefit (Nam et al. (2017)). To study effective modulation of exoskeletal assistance strategies, we modeled an exoskeleton for the elbow using an ideal actuator and the addition of two supports with a weight of 0.101 Kg for the upper arm and 0.111 Kg on the forearm (Figure 1-D, Wang et al. (2022)). The assistance given by the exoskeleton was a percentage of the biological joint torque, this was based on the neuromusculoskeletal controller presented in Durandau et al. (2019).

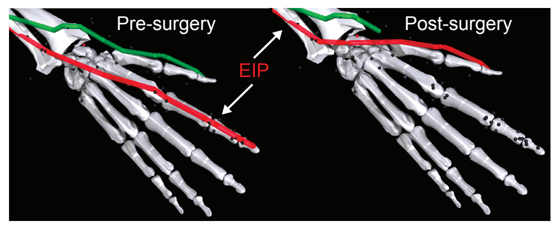

Figure 3: Tendon transfer. Left: EIP (in red - controls the Index extensor) and EPL (in green - controls the thumb extensor) muscles on a healthy hand. Right: the muscle path of EIP is re-routed on the EPL as thumb extensor.

4. Experiments

4.1. Models validation

We compared the MuJoCo models against the correspondent OpenSim models. Muscle moment arm root mean square (RMS) differences between the MuJoCo model with respect to the Opensim model were $0.044 \pm 0.09\%$ for the elbow and $0.38 \pm 0.57\%$ for the hand model. Also, the RMS error in forces was $2.2 \pm 1.4\%$ F_{max} (OpenSim peak force) for the elbow model and $4.1 \pm 2.0\%$ F_{max} for the hand model. Those errors indicate that the MuJoCo models are anatomically and dynamically similar to the SOTA OpenSim model. Forward simulations showed that MuJoCo models can be several orders of magnitude faster than OpenSim (see Figure 4, from 70x to 4000x faster). By simulating an elbow model (6 muscles) where we iteratively replicated all muscles, it was possible to observe that the OpenSim computing time increased exponentially while the MuJoCo did not (see Figure 4). This increase in efficiency is mostly the result of a simplified implementation of the muscle actuator in

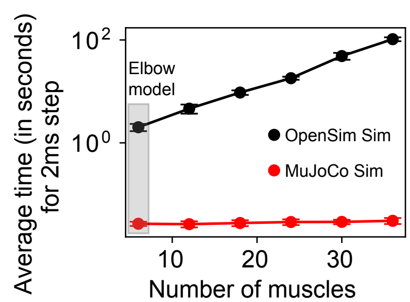

Figure 4: Comparison between MuJoCo and OpenSim Models speed when the same muscles are proportionally replicated.

MuJoCo, which allows faster and more stable simulations. In summary, training/studies (Section 4.2) that in OpenSim will take years to simulate can be performed in MuJoCo in a few days!

4.2. Baselines

We present baseline results obtained using Natural Policy Gradient (NPG, Kakade (2001)) for a subset (i.e. stationary, easy and hard, with fixed resets) of the conditions available in MyoSuite. We chose this algorithm owing to its recent successes in solving complex robotic dexterous manipulation tasks (Rajeswaran et al. (2018b)). In Figure 5, we show success rates for the different tasks up to 5M samples. More complex tasks e.g. Baoding balls can be solved, albeit with much higher sample complexity (∼70M samples). In Figure 6A, it is shown a sequence of snapshots of the solution of the key turning task. It is possible to see how the index and thumb activation is functional to the effective rotation of the key. The Pen Twirl task (Figure 6B) requires effective coordination between the wrist and the hand muscles to express the full dexterity of the hand while effectively maneuvering the pen (blue) to the desired goal (green) via a sequences of contacts. Finally, the task of baoding balls (Figure 6C) pushes the dexterity of the hand to its limits by requiring policies to learn simultaneous coordination of not one but two objects. This task is quite challenging to learn as any miss coordination results in catastrophic failures.

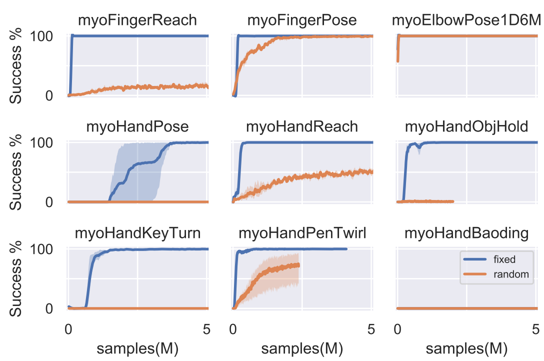

Figure 5: Tasks performance in the easy (fixed) and the hard (random) conditions with fixed reset.

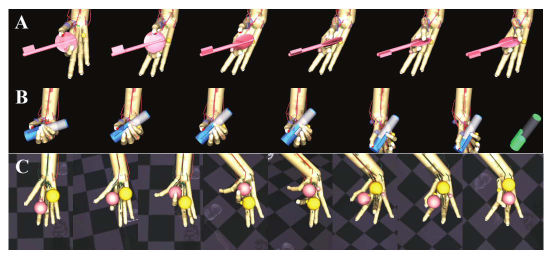

Figure 6: Frames depicting complex wrist and finger coordination behaviors learned to solve the Key Turn (A), the Pen Twirl (B), and the Baoding Balls (C) tasks.

4.3. Intrinsic non-stationary perturbations

While extrinsic perturbation can be easily added to each task, we focus our next investigation on behavioral acquisition and adaptation in response to intrinsic non-stationary perturbations (Sec 3.3) that are modelled after various real-life scenarios in MyoSuite.

4.3.1. SARCOPENIA (MUSCLE DEGENERATION)

We tested a policy trained on a 1D elbow flexion task to reach random targets in its operational space on alternated movements between points A and B (see Figure 7A).

First, we study how sarcopenia (muscle weakness) affects the control of movement. In presence of Sarcopenia (Figure 7B) the Brachioradialis (BRA) - which in normal conditions does not need support from other muscles - needs stronger activations from synergistic muscles (BICLong - biceps longus - and BICShort - biceps short) to solve the task.

4.3.2. FATIGUE (MUSCLE EXHAUSTION)

Second, we investigated the effect of muscle fatigue (introduced in Section 3.3). In this case, the loss of muscle power is progressive over time. In the same alternated movements elbow tasks, we observe gradually increasing contribution of synergistic muscles to compensate for the muscle force loss (Figure 7C).

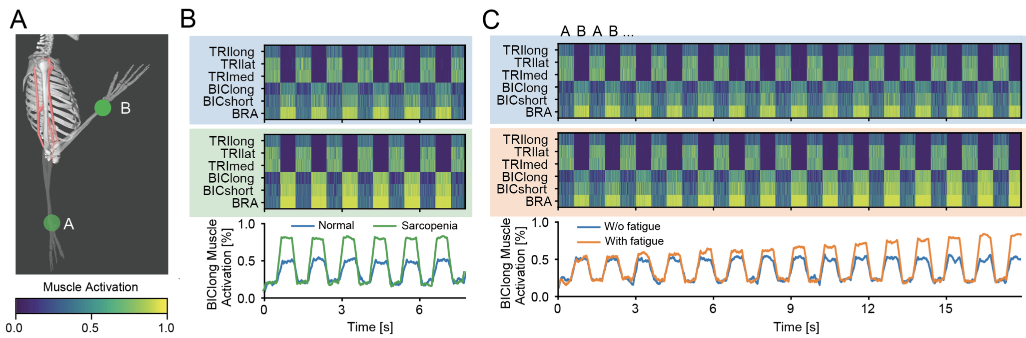

Figure 7: Control Policy with muscle sarcopenia (weakness) and fatigue during an alternated reaching task between two targets. (A) Experiment design. (B) Effects of Sarcopenia – Top-row: muscle activations for 6 arm muscles (3 agonists and 3 antagonists). Middle-row: muscle activations in presence of sarcopenia. All synergistic flexor muscles (BRA - brachioradialis, BICLong - Biceps Long, BICShort - Biceps short) increase their activation to compensate for the reduced force. Bottom-row: a trace of the activation only for the Biceps Long muscle. (C) Effects of Fatigue – Bottom-row: synergistic muscles progressively increase activations to compensate for Fatigue.

4.3.3. ACCIDENTS (TENDON TEAR)

Next, in order to study the effects of tendon tear, a policy was first trained to solve the Key-Turn task and then challenged with the selective damage(tear) of different thumb muscles (Figure 8). While there are redundancies (not every tear crippled ability), we found that FPL and OP are critical for the correct solution of the key task. In absence of OP the FPL is able to compensate. But when OP is torn FPL cannot compensate resulting in reduced key rotations.

4.3.4. CONTROL IN PRESENCE OF TENDON TRANSFER

Finally, a tendon transfer - EIP muscle was routed to replace the EPL muscle - was performed to test the ability of the policies to compensate for action re-mapping due to tendon surgery. After the surgery, the major thumb muscles needed to compensate for the different activation space (Figure 9 and, a previously trained policy was unable to solve the task. Indeed it was necessary an extensive additional training to control the thumb after tendon transfer (Online Appendix/Solutions). This is typically observed in patient undergoing extensive physiotherapeurical sessions to re-learn the control of the thumb (Wangdell et al. (2016)).

5. Discussion and Conclusions

Here, we have proposed a new framework of physiologically realistic and computationally efficient models and tasks to study human motor control. The proposed models include highly skilled ma-

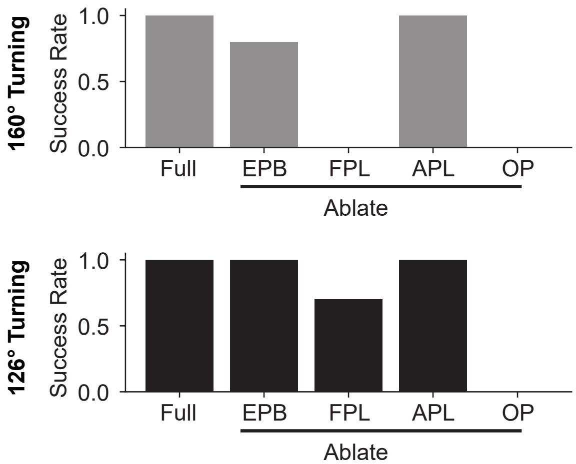

Figure 8: Effects of tendon-tear on key turning task. After training, selectively a different thumb muscle was removed for each experiment: 2 flexors (Flexor Pollicis Longus (FPL), Opponens Pollicis (OP)) and 1 extensor (Extensor Pollicis Brevis (EPB)) and 1 abductor (Abductor Pollicis Brevis (APL)), and the task was repeated 10 times. Success rates in turning the key to reach a threshold of 126° or 160° rotation.

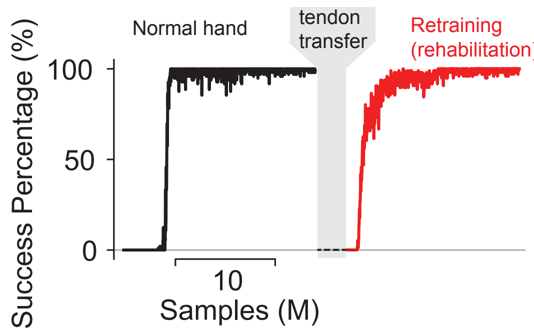

Figure 9: Tendon transfer on key turning task. In this experiment we first test a key-turning task on an intact hand and then we operated a tendon transfer: activity of the muscle Extensor Indicis Proprius (EIP) was redirected to activate muscle Extensor Pollicis Longus (EPL). After tendon transfer, the policy needs to be retrained in order to solve the task.

nipulations and realistic non-stationarities to simulate real-life scenarios such as muscle fatigue, sarcopenia, and tendon transfer. This benchmark will provide biologically relevant problems where task success and physiological representations might differ. Actually, because activation of the musculoskeletal models can be directly related to a (normalized) intramuscular activity recorded in human subjects, it will be possible to validate experimentally the learned solutions.

The basic policies trained already showed physiologically relevant behaviors. First, we observe an automatic adaptation of muscles to leverage antagonistic effects of flexors and extensors. This was clearly visible with the elbow model where those effects could be more easily isolated. Second, co-contractions evolve naturally to compensate for changes in muscle properties e.g. sarcopenia and fatigue. Nevertheless, those compensations are not enough to replace muscles like the Opponens Pollicis which has unique functions for hand manipulations.

Importantly, the implemented models are only the initial iteration on approximating the musculoskeletal system that will need further development and validation. Indeed, both the tasks and physiological changes in muscle properties are a subset of the possible changes that can be considered.

All in all, this work can facilitate cross-pollination and catalyzation of new ideas between different communities. The ML community will have more challenging tasks with 3rd order dynamics (uncommon in ML/RL benchmarks) and physiological realistic non-stationarity provides. The biomechanics community will have an additional platform where contact-rich interactions at scale can be studied. Finally, the robotic community will have a biological system to develop strategies for. Overall, this work will allow more realistic simulation allowing *in sillico* trials of humans, robots, and their interaction.

Acknowledgments

H.W., G.D., and M.S. received funding by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme as part of the ERC Starting Grant INTERACT (Grant No. 803035) as well as part by the Horizon 2020 ICT-10 Project SOPHIA (871237).

References

Simm: Software for interactive musculoskeletal modeling. https://motionanalysis.com/simm/.

Marko Ackermann and Antonie J Van den Bogert. Optimality principles for model-based prediction of human gait. *Journal of biomechanics*, 43(6):1055–1060, 2010.

OpenAI: Marcin Andrychowicz, Bowen Baker, Maciek Chociej, Rafal Jozefowicz, Bob McGrew, Jakub Pachocki, Arthur Petron, Matthias Plappert, Glenn Powell, Alex Ray, et al. Learning dexterous in-hand manipulation. *The International Journal of Robotics Research*, 39(1):3–20, 2020.

Edith M. Arnold and Scott L. Delp. Fibre operating lengths of human lower limb muscles during walking. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366:1530 – 1539, 2011.

Nikolai Bernstein. The co-ordination and regulation of movements. *The co-ordination and regulation of movements*, 1966.

Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. Openai gym. *arXiv preprint arXiv:1606.01540*, 2016.

Don B Chaffin, Gunnar B J Andersson, and Bernard J Martin. *Occupational biomechanics*. John wiley \& sons, 2006.

Tao Chen, Jie Xu, and Pulkit Agrawal. A system for general in-hand object re-orientation. *arXiv preprint arXiv:2111.03043*, 2021.

Erwin Coumans and Yunfei Bai. Pybullet, a python module for physics simulation for games, robotics and machine learning. 2016.

Alfonso J. Cruz-Jentoft and Avan A. Sayer. Sarcopenia. *The Lancet*, 393(10191):2636–2646, 6 2019. ISSN 0140-6736. doi: 10.1016/S0140-6736(19)31138-9.

Michael Damsgaard, John Rasmussen, Søren Tørholm Christensen, Egidijus Surma, and Mark De Zee. Analysis of musculoskeletal systems in the anybody modeling system. *Simulation Modelling Practice and Theory*, 14(8):1100–1111, 2006.

Scott L. Delp, Frank C. Anderson, Allison S. Arnold, Peter Loan, Ayman Habib, Chand T. John, Eran Guendelman, and Darryl G. Thelen. Opensim: Open-source software to create and analyze dynamic simulations of movement. *IEEE Transactions on Biomedical Engineering*, 54(11): 1940–1950, 2007. doi: 10.1109/TBME.2007.901024.

Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009. doi: 10.1109/CVPR.2009.5206848.

Coline Devin, Abhishek Gupta, Trevor Darrell, Pieter Abbeel, and Sergey Levine. Learning modular neural network policies for multi-task and multi-robot transfer. In *2017 IEEE international conference on robotics and automation (ICRA)*, pages 2169–2176. IEEE, 2017.

Richard M. Dodds, Holly E. Syddall, Rachel Cooper, Diana Kuh, Cyrus Cooper, and Avan Aihie Sayer. Global variation in grip strength: a systematic review and meta-analysis of normative data. *Age and Ageing*, 45(2):209–216, 3 2016. ISSN 0002-0729. doi: 10.1093/AGEING/AFV192. URL https://academic.oup.com/ageing/article/45/2/209/2195303.

Guillaume Durandau, Dario Farina, Guillermo Asín-Prieto, Iris Dimbwadyo-Terrer, Sergio Lerma-Lara, Jose L Pons, Juan C Moreno, and Massimo Sartori. Voluntary control of wearable robotic exoskeletons by patients with paresis via neuromechanical modeling. *Journal of neuroengineering and rehabilitation*, 16(1):1–18, 2019.

Tom Erez, Yuval Tassa, and Emanuel Todorov. Simulation tools for model-based robotics: Comparison of bullet, havok, mujoco, ODE and physx. In *IEEE International Conference on Robotics and Automation, ICRA 2015*, pages 4397–4404. IEEE, 2015. doi: 10.1109/ICRA.2015.7139807. URL https://doi.org/10.1109/ICRA.2015.7139807.

R. I. Gelb. Tendon Transfer for Rupture of the Extensor Pollicis Longus. *Hand Clinics*, 11(3): 411–422, 8 1995. ISSN 0749-0712. doi: 10.1016/S0749-0712(21)00062-7.

Friedl De Groote, Allison L. Kinney, Anil V. Rao, and Benjamin J. Fregly. Evaluation of direct collocation optimal control problem formulations for solving the muscle redundancy problem. *Annals of Biomedical Engineering*, 44:2922 – 2936, 2016.

Abhishek Gupta, Coline Devin, YuXuan Liu, Pieter Abbeel, and Sergey Levine. Learning invariant feature spaces to transfer skills with reinforcement learning. *arXiv preprint arXiv:1703.02949*, 2017.

Abhishek Gupta, Justin Yu, Tony Z Zhao, Vikash Kumar, Aaron Rovinsky, Kelvin Xu, Thomas Devlin, and Sergey Levine. Reset-free reinforcement learning via multi-task learning: Learning dexterous manipulation behaviors without human intervention. *arXiv preprint arXiv:2104.11203*, 2021.

Samuel R Hamner, Ajay Seth, and Scott L Delp. Muscle contributions to propulsion and support during running. *Journal of biomechanics*, 43(14):2709–2716, 2010.

Masaya Hirashima and Tomomichi Oya. How does the brain solve muscle redundancy? filling the gap between optimization and muscle synergy hypotheses. *Neuroscience Research*, 104:80–87, 2016. ISSN 0168-0102. doi: https://doi.org/10.1016/j.neures.2015.12.008. URL https://www.sciencedirect.com/science/article/pii/S0168010215002990. Body representation in the brain.

Wenlong Huang, Igor Mordatch, Pieter Abbeel, and Deepak Pathak. Generalization in dexterous manipulation via geometry-aware multi-task learning. *arXiv preprint arXiv:2111.03062*, 2021.

Jemin Hwangbo, Joonho Lee, and Marco Hutter. Per-contact iteration method for solving contact dynamics. *IEEE Robotics and Automation Letters*, 3(2):895–902, 2018. URL www.raisim.com.

Aleksi Ikkala and Perttu Hämäläinen. Converting biomechanical models from opensim to mujoco, 2020.

Sašo Jezernik, Gery Colombo, Thierry Keller, Hansruedi Frueh, and Manfred Morari. Robotic orthosis lokomat: A rehabilitation and research tool. *Neuromodulation: Technology at the neural interface*, 6(2):108–115, 2003.

Sham M Kakade. A natural policy gradient. *Advances in neural information processing systems*, 14, 2001.

Fotios Alexandros Karakostis, Daniel Haeufle, Ioanna Anastopoulou, Konstantinos Moraitis, Gerhard Hotz, Vangelis Tourloukis, and Katerina Harvati. Biomechanics of the human thumb and the evolution of dexterity. *Current Biology*, 31(6):1317–1325.e8, 2021. ISSN 0960-9822. doi: https://doi.org/10.1016/j.cub.2020.12.041. URL https://www.sciencedirect.com/science/article/pii/S0960982220318935.

Donald T. Kirkendall and William E. Garrett. The effects of aging and training on skeletal muscle. *The American Journal of Sports Medicine*, 26(4):598–602, 1998. doi: 10.1177/03635465980260042401. URL https://doi.org/10.1177/03635465980260042401. PMID: 9689386.

Vijay R Konda and John N Tsitsiklis. Actor-critic algorithms. In *Advances in neural information processing systems*, pages 1008–1014, 2000.

Mark L Latash. The bliss (not the problem) of motor abundance (not redundancy). *Experimental brain research*, 217(1):1–5, 2012.

Jeongseok Lee, Michael X Grey, Sehoon Ha, Tobias Kunz, Sumit Jain, Yuting Ye, Siddhartha S Srinivasa, Mike Stilman, and C Karen Liu. Dart: Dynamic animation and robotics toolkit. *Journal of Open Source Software*, 3(22):500, 2018.

Jong Hwa Lee, Deanna S. Asakawa, Jack T. Dennerlein, and Devin L. Jindrich. Finger muscle attachments for an opensim upper-extremity model. *PLOS ONE*, 10(4):1–28, 04 2015. doi: 10.1371/journal.pone.0121712. URL https://doi.org/10.1371/journal.pone.0121712.

Seunghwan Lee, Moonseok Park, Kyoungmin Lee, and Jehee Lee. Scalable muscle-actuated human simulation and control. *ACM Trans. Graph.*, 38(4), jul 2019. ISSN 0730-0301. doi: 10.1145/3306346.3322972. URL https://doi.org/10.1145/3306346.3322972.

Liang Ma, Damien Chablat, Fouad Bennis, and Wei Zhang. A new simple dynamic muscle fatigue model and its validation. *International Journal of Industrial Ergonomics*, 39(1):211–220, 2009. ISSN 01698141. doi: 10.1016/j.ergon.2008.04.004. URL http://dx.doi.org/10.1016/j.ergon.2008.04.004.

Viktor Makoviychuk, Lukasz Wawrzyniak, Yunrong Guo, Michelle Lu, Kier Storey, Miles Macklin, David Hoeller, Nikita Rudin, Arthur Allshire, Ankur Handa, and Gavriel State. Isaac gym: High performance GPU based physics simulation for robot learning. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*, 2021. URL https://openreview.net/forum?id=fgFBtYgJQX_.

Daniel C. McFarland, Emily M. McCain, Michael N. Poppo, and Katherine R. Saul. Spatial dependency of glenohumeral joint stability during dynamic unimanual and bimanual pushing and pulling. *Journal of Biomechanical Engineering-Tansation of the ASME*, 141(5), 05 2019. doi: 10.1115/1.4043035.

Anusha Nagabandi, Ignasi Clavera, Simin Liu, Ronald S Fearing, Pieter Abbeel, Sergey Levine, and Chelsea Finn. Learning to adapt in dynamic, real-world environments through meta-reinforcement learning. *arXiv preprint arXiv:1803.11347*, 2018.

Anusha Nagabandi, Kurt Konolige, Sergey Levine, and Vikash Kumar. Deep dynamics models for learning dexterous manipulation. In *Conference on Robot Learning*, pages 1101–1112. PMLR, 2020.

Ki Yeun Nam, Hyun Jung Kim, Bum Sun Kwon, Jin-Woo Park, Ho Jun Lee, and Aeri Yoo. Robot-assisted gait training (lokomat) improves walking function and activity in people with spinal cord injury: a systematic review. *Journal of neuroengineering and rehabilitation*, 14(1):1–13, 2017.

OpenAI, Ilge Akkaya, Marcin Andrychowicz, Maciek Chociej, Mateusz Litwin, Bob McGrew, Arthur Petron, Alex Paino, Matthias Plappert, Glenn Powell, Raphael Ribas, Jonas Schneider, Nikolas Tezak, Jerry Tworek, Peter Welinder, Lilian Weng, Qiming Yuan, Wojciech Zaremba, and Lei Zhang. Solving rubik's cube with a robot hand. *arXiv preprint arXiv:1910.07113*, 2019.

Vassil Panayotov, Guoguo Chen, Daniel Povey, and Sanjeev Khudanpur. Librispeech: An asr corpus based on public domain audio books. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5206–5210, 2015. doi: 10.1109/ICASSP.2015.7178964.

Matthias Plappert, Marcin Andrychowicz, Alex Ray, Bob McGrew, Bowen Baker, Glenn Powell, Jonas Schneider, Josh Tobin, Maciek Chociej, Peter Welinder, Vikash Kumar, and Wojciech Zaremba. Multi-goal reinforcement learning: Challenging robotics environments and request for research, 2018.

Aravind Rajeswaran, Vikash Kumar, Abhishek Gupta, Giulia Vezzani, John Schulman, Emanuel Todorov, and Sergey Levine. Learning Complex Dexterous Manipulation with Deep Reinforcement Learning and Demonstrations. In *Proceedings of Robotics: Science and Systems (RSS)*, 2018a.

Aravind Rajeswaran, Vikash Kumar, Abhishek Gupta, Giulia Vezzani, John Schulman, Emanuel Todorov, and Sergey Levine. Learning Complex Dexterous Manipulation with Deep Reinforcement Learning and Demonstrations. In *Proceedings of Robotics: Science and Systems (RSS)*, 2018b.

Massimo Sartori, Leonardo Gizzi, David G Lloyd, and Dario Farina. A musculoskeletal model of human locomotion driven by a low dimensional set of impulsive excitation primitives. *Frontiers in computational neuroscience*, 7:79, 2013.

Katherine R. Saul, Xiao Hu, Craig M. Goehler, Meghan E. Vidt, Melissa Daly, Anca Velisar, and Wendy M. Murray. Benchmarking of dynamic simulation predictions in two software platforms using an upper limb musculoskeletal model. *Computer Methods in Biomechanics and Biomedical Engineering*, 18(13):1445–1458, 2015a. doi: 10.1080/10255842.2014.916698. URL https://doi.org/10.1080/10255842.2014.916698. PMID: 24995410.

Katherine R Saul, Xiao Hu, Craig M Goehler, Meghan E Vidt, Melissa Daly, Anca Velisar, and Wendy M Murray. Benchmarking of dynamic simulation predictions in two software platforms using an upper limb musculoskeletal model. *Computer methods in biomechanics and biomedical engineering*, 18(13):1445–1458, 2015b.

Stephen H. Scott. The computational and neural basis of voluntary motor control and planning. *Trends in Cognitive Sciences*, 16(11):541–549, 2012. ISSN 1364-6613. doi: https://doi.org/10.1016/j.tics.2012.09.008. URL https://www.sciencedirect.com/science/article/pii/S1364661312002240.

Ajay Seth, Jennifer L. Hicks, Thomas K. Uchida, Ayman Habib, Christopher L. Dembia, James J. Dunne, Carmichael F. Ong, Matthew S. DeMers, Apoorva Rajagopal, Matthew Millard, Samuel R. Hamner, Edith M. Arnold, Jennifer R. Yong, Shrinidhi K. Lakshmikanth, Michael A. Sherman, Joy P. Ku, and Scott L. Delp. Opensim: Simulating musculoskeletal dynamics and neuromuscular control to study human and animal movement. *PLOS Computational Biology*, 14: 1–20, 07 2018. doi: 10.1371/journal.pcbi.1006223. URL https://doi.org/10.1371/journal.pcbi.1006223.

Reza Shadmehr and John W Krakauer. A computational neuroanatomy for motor control. *Experimental brain research*, Mar;185(3):359-81, 2008.

David Silver, Guy Lever, Nicolas Heess, Thomas Degris, Daan Wierstra, and Martin Riedmiller. Deterministic policy gradient algorithms. In *International conference on machine learning*, pages 387–395. PMLR, 2014.

Seungmoon Song, Łukasz Kidziński, Xue Bin Peng, Carmichael Ong, Jennifer Hicks, Sergey Levine, Christopher G. Atkeson, and Scott L. Delp. Deep reinforcement learning for modeling human locomotion control in neuromechanical simulation. *bioRxiv*, 2020. doi: 10.1101/2020.08.11.246801.

Richard S Sutton. Generalization in reinforcement learning: Successful examples using sparse coarse coding. *Advances in neural information processing systems*, pages 1038–1044, 1996.

Yuval Tassa, Yotam Doron, Alistair Muldal, Tom Erez, Yazhe Li, Diego de Las Casas, David Budden, Abbas Abdolmaleki, Josh Merel, Andrew Lefrancq, et al. Deepmind control suite. *arXiv preprint arXiv:1801.00690*, 2018.

Lena H. Ting, Stacie A. Chvatal, Seyed A. Safavynia, and J. Lucas McKay. Review and perspective: neuromechanical considerations for predicting muscle activation patterns for movement.

International Journal for Numerical Methods in Biomedical Engineering, 28(10):1003–1014, 2012. doi: https://doi.org/10.1002/cnm.2485. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.2485.

Emanuel Todorov, Tom Erez, and Yuval Tassa. Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5026–5033. IEEE, 2012.

Nina K Vøllestad. Measurement of human muscle fatigue. *Journal of neuroscience methods*, 74(2): 219–227, 1997.

Alex Wang, Amanpreet Singh, Julian Michael, Felix Hill, Omer Levy, and Samuel R. Bowman. GLUE: A multi-task benchmark and analysis platform for natural language understanding. *CoRR*, abs/1804.07461, 2018. URL http://arxiv.org/abs/1804.07461.

Huawei* Wang, Vittorio* Caggiano, Guillaume Durandau, Kumar Sartori, Massimo, and Vikash. Myosim: Fast and physiologically realistic mujoco models for musculoskeletal and exoskeletal studies. In *2022 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2022.

Johanna Wangdell, Lina Bunketorp-Käll, Sabrina Koch-Borner, and Jan Fridén. Early active rehabilitation after grip reconstructive surgery in tetraplegia. *Archives of physical medicine and rehabilitation*, 97(6):S117–S125, 2016.

Christopher JCH Watkins and Peter Dayan. Q-learning. *Machine learning*, 8(3-4):279–292, 1992.

Scott C White, H John Yack, and David A Winter. A three-dimensional musculoskeletal model for gait analysis. anatomical variability estimates. *Journal of biomechanics*, 22(8-9):885–893, 1989.

Daniel M. Wolpert and Zoubin Ghahramani. Computational principles of movement neuroscience. *Nature Neuroscience*, 3 Suppl 1:1212–1217, 2000.

John Z Wu, Kai-Nan An, Robert G Cutlip, Kristine Krajnak, Daniel Welcome, and Ren G Dong. Analysis of musculoskeletal loading in an index finger during tapping. *Journal of biomechanics*, 41(3):668–676, 2008.

Zhe Xu, Vikash Kumar, Yoky Matsuoka, and Emanuel Todorov. Design of an anthropomorphic robotic finger system with biomimetic artificial joints. In *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, pages 568–574. IEEE, 2012.

Kim Younguk, Yihwan Jung, Woosung Choi, Kunwoo Lee, and Seungbum Koo. Similarities and differences between musculoskeletal simulations of opensim and anybody modeling system. *Journal of Mechanical Science and Technology*, 32:6037–6044, 12 2018. doi: 10.1007/s12206-018-1154-0.