



HAL
open science

Fatigue Mitigation through Planning in Human-Robot Repetitive Co-Manipulation: Automatic Extraction of Relevant Action Sets

Aya Yaacoub, Vincent Thomas, Francis Colas, Pauline Maurice

► **To cite this version:**

Aya Yaacoub, Vincent Thomas, Francis Colas, Pauline Maurice. Fatigue Mitigation through Planning in Human-Robot Repetitive Co-Manipulation: Automatic Extraction of Relevant Action Sets. 2024 IEEE-RAS 23rd International Conference on Humanoid Robots (Humanoids), Nov 2024, Nancy, France. pp.795-802, 10.1109/Humanoids58906.2024.10769843 . hal-04830173

HAL Id: hal-04830173

<https://hal.science/hal-04830173v1>

Submitted on 10 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Fatigue Mitigation through Planning in Human-Robot Repetitive Co-Manipulation: Automatic Extraction of Relevant Action Sets

Aya Yaacoub¹, Vincent Thomas¹, Francis Colas¹, Pauline Maurice¹

Abstract— Work-related musculoskeletal disorders (WMSDs) are among the most common injuries associated with industrial tasks. Repetitive tasks are a major WMSDs risk factor, because they load the same human joints over and over again. Collaborative robots can be used to induce movement variability in highly repetitive co-manipulation tasks by changing the position of the co-manipulated object through time, thereby distributing the physical load over different body parts and reducing fatigue accumulation. This is even more beneficial when long-term consequences of the robot actions are considered. However, selecting the optimal action within the continuous robot workspace is not compatible with time constraints imposed by online planning in highly repetitive tasks, especially when the planning horizon increases. In this work we therefore propose an approach to automatically extract a set of actions from the continuous workspace, that combines two properties: planning speed (i.e. reduced number of actions in the set), and ability to induce a variety of fatigue distributions over the different human joints. The proposed approach combines a digital human simulation to estimate the fatigue induced by possible actions, with a repeated short-term planning (greedy-based selection approach) phase that explores the fatigue space and simultaneously identifies optimal actions from a large space for each visited state. By retaining actions used in the short-term planning, this process allows to extract a subset of relevant actions. We evaluate our approach in a simulated co-manipulation scenario, and show that the resulting action set robustly outperforms action sets extracted with benchmark methods, both in terms of planning time and human fatigue mitigation.

I. INTRODUCTION

Work-related musculoskeletal disorders (WMSDs) are the most common occupational disease in many countries, and reach a high prevalence among industrial workers [1], [2]. Repetitive manual tasks have been identified as a major WMSDs risk factor, even with limited force exertion, because mechanical stress repeatedly affects the same musculoskeletal structures [3]. Actually, studies in the field of ergonomics suggest that movement variability, i.e. varying the posture adopted to perform a task, could be beneficial to reduce WMSDs since it distributes mechanical load over different body joints over time [4].

With the promise of collaborative robots (cobots) directly interacting with humans, a research direction investigating using cobots to reduce WMSDs risk has recently emerged [5], [6]. In particular, several studies proposed to use cobots to induce changes in the human posture during repetitive co-manipulation tasks, using reactive robot behavior [7], random

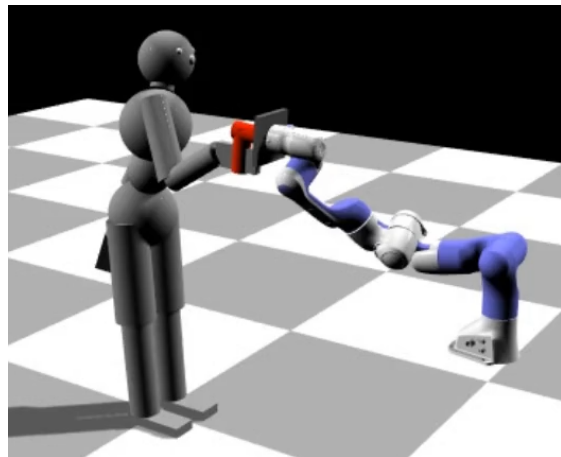


Fig. 1. Screenshot of the physics-based digital human simulation used to estimate the fatigue evolution induced by an action (i.e. end-effector pose) of the collaborative robot. The simulation serves to compute the whole-body motion of the human in reaction to the robot action, and the associated human joint torques. A fatigue model is then used to derive fatigue evolution from the joint torques.

changes [8], or long-term planning [9]. By modifying the human posture, these approaches aim at distributing the physical stress on different body parts over time, hence reducing fatigue effect. In a previous work, we introduced a probabilistic planning approach to optimize a cobot behavior, in order to reduce human physical fatigue (used as a proxy for WMSDs risk) in a highly repetitive task [9]. A specificity of this approach was to take into account the inherent uncertainties on the human internal state and future posture, stemming from the human kinematic redundancy. Importantly, we showed that planning over several task cycles was beneficial compared to reasoning over a single cycle[9].

However, the complexity of probabilistic planning –and hence the decision time– increases dramatically with both the planning horizon and the size of the action space [10]. When planning the robot behavior in human-robot co-manipulation, actions correspond to the robot end-effector poses or trajectories (depending on the specific task), hence the action space is continuous. Methods have been proposed to deal with planning in a continuous action space, mainly through online sampling of the action space. For instance, Progressive Widening (PW) algorithms progressively add actions to an initial action set alongside planning [11][12]. However, an action set is refined specifically for one belief state of the system, hence the (potentially time-consuming) process needs

*This work was partly supported by the French ANR under Grant No. ANR-20-CE33-0004 (project ROOIBOS).

¹ The authors are with Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France. `firstname.name@loria.fr`

to be repeated for each new belief, i.e. at each new step of the planning process. Such online-sampling approaches are therefore hardly compatible with time constraints associated with online robot decision-making, especially in the context of highly repetitive tasks where cycle time varies from a few seconds to a few minutes.

This work therefore addresses the question of offline automatic selection of a reduced set of robot actions relevant to long-term human fatigue mitigation. The challenge lies in the tradeoff between the ability to induce a large variety of human postures to distribute fatigue over the whole body (for which more actions is better), and the speed of planning (for which less actions is better). Specifically, working on fatigue distribution requires to consider fatigue as a high-dimensional state (each component corresponding to the fatigue of a specific muscle or joint). Hence the relevance of an action depends on its complementarity with the other actions in the set, in terms of diversity of consequences in a high-dimensional space.

We propose an offline approach for action-set extraction that combines physics-based digital human simulation and short-term planning. A 1-step horizon planning is repeatedly run to select the (short-term) optimal action within the large action space, using digital human simulation to evaluate the fatigue evolution induced by all possible actions. The process is repeated over a large number of cycles and in different initial conditions to allow a wide exploration of the fatigue space, and hence the selection of actions relevant for a diversity of situations. We test our approach in simulation to benchmark the long-term planning performance of the resulting action set in terms of online decision time and fatigue mitigation ability.

II. METHODS

The objective of our approach is to extract a set of actions from a continuous action space, which represents a good compromise between online decision-making time (i.e. minimal number of actions in the set) and ability to distribute fatigue on different body parts over time. In the remaining of this article, we illustrate our approach on a task where, at each cycle, a collaborative robot brings an object to the human co-worker at a certain Cartesian pose (Fig. 1, video ¹). The human then reaches to the object to work on it while the robot continues to hold it at the same pose (e.g. spray painting car pieces, scanning parcels packed by the robot). When done, the robot moves the object away and brings a new one, which corresponds to starting a new cycle. An action therefore corresponds to a robot end-effector Cartesian pose, and the continuous action space is the intersection of the human and robot workspaces.

This section first presents the formulation of the long-term decision making problem (Section II-A), and how it can be solved (Section II-B). Section II-C then describes the digital human simulation used to estimate the fatigue induced by a robot action. Finally, the proposed approach for action-set extraction is detailed in Section II-D.

¹<https://youtu.be/mBChsgYRX1g>

A. Fatigue Mitigation Problem Formulation

The objective of the robot in the considered scenario is to induce postural changes in the human by varying the co-manipulated object pose across cycles, in order to minimize the accumulated fatigue over a long-term horizon (i.e. many cycles). However, owing to human kinematic redundancy, the human postural reaction (i.e. whole-body motion) is not fully determined by the object pose: the human may adopt different postures to perform a same task, resulting in different fatigue evolutions. This uncertainty on the human reaction therefore needs to be taken into account when selecting the optimal robot action. In addition, the reasoning is done on the human physical fatigue, which is an internal state that cannot directly be measured but only inferred. The problem is then only partially observable. For the above-listed reasons (long-term planning, uncertainty, partial observability), we model the problem using the Partially Observable Markov Decision Process (POMDP) framework [13].

The POMDP describes possible transitions from a partially-observable human fatigue state $s \in S$, when a robot action $a \in A$ is executed, to some new fatigue state $s' \in S$ with a probability $T(s, a, s')$. Following the transition (s, a, s') , an observation $o \in \Omega$ of the human postural reaction is returned with a probability $O(s', o)$, which serves to update the robot belief on the current human fatigue state. A reward $r = R(s')$ is associated with each state s' and depends on the fatigue value at s' . Finally, the initial belief state $b_0 \in B$ is a probability distribution representing the uncertainty on the initial human fatigue state (i.e. at the beginning of the first cycle).

B. Robot Policy Computation

Solving the POMDP consists in finding a policy $\pi : B \mapsto A$ according to which the next robot action a will be selected at each cycle, given the current belief b on the human fatigue state. The optimal policy π^* maximizes the average discounted-sum of received rewards $\mathbb{E}[\sum_{t=0}^{\infty} \gamma^t r_t]$ during execution, with γ being the discount factor balancing the importance of immediate vs. late rewards. The optimal policy can be approximated using algorithms such as POMCP based on Monte-Carlo Tree Search [14].

However, in situations where decision time is critical, a finite-horizon algorithm can be a faster sub-optimal alternative to solve the POMDP. A finite horizon algorithm guarantees a balanced evaluation of all possible actions within an estimable time controlled by the choice of the planning horizon. The principle of the finite-horizon (FH) algorithm is summarized in Algorithm 1. For a given horizon $h \in \mathbb{N}$, and starting from a belief b over the human fatigue state, FH_h returns the action leading to the highest expected sum of discounted rewards after h cycles².

C. Fatigue Evolution Estimation

Computing the robot policy requires to evaluate the rewards associated with a series of actions. Since the reward

²Note that FH_1 is equivalent to *greedy* action selection, i.e. selecting the action that minimizes the expected fatigue cost after one single cycle.

Algorithm 1 Finite Horizon FH

Input h, b_0 \triangleright planning horizon, initial belief
Output a_{max} \triangleright action with highest value
 $a_{max}, V_{a_{max}} \leftarrow FH(h, b_0)$
return a_{max}

procedure $FH(h, b)$
for $a \in A$ **do**
 Initialize $V_a \leftarrow 0$ \triangleright value of action a
 for $o \in \Omega$ **do**
 $b'(s')_{s' \in S} \propto O(s', o) \sum_{s \in S} T(s, a, s').b(s)$
 $a', V_{a'} \leftarrow FH(h-1, b')$ \triangleright recursive call
 $V_a^o \leftarrow r(b, a, o) + V_{a'}$
 $V_a \leftarrow V_a + P(o|b, a) \cdot V_a^o$
 return $tuple(\arg \max_{a \in A} V_a, \max_{a \in A} V_a)$

depends here on the fatigue state, it means predicting the fatigue evolution induced by the series of actions.

In this work, we consider fatigue at joint level, and define it as a temporary loss in the torque generation capacity of a joint j [15]:

$$f^j(t) = 1 - \frac{\tau_{cem}^j(t)}{\tau_{max}^j} \quad (1)$$

where f^j is the instantaneous fatigue of joint j , τ_{cem}^j is the current maximum exertable torque of joint j , and τ_{max}^j is the nominal maximum exertable torque (in the absence of fatigue)³. We use the model of Ma *et al.* to represent the evolution of $\tau_{cem}^j(t)$ as a function of the history of joint torques $\tau_{0 \rightarrow t}^j$ exerted during the motion [16].

Estimating the non-measurable fatigue evolution induced by a robot action therefore requires to estimate the human joint torques exerted during the human postural reaction (whole-body motion) in response to the robot action. In this work, we use a physics-based Digital Human Model (DHM) simulation (Fig. 1) to generate dynamically-consistent whole-body motions and their associated joint torques from a high-level description of the task to perform (defined by the pose of the co-manipulated object). The DHM motion is computed by solving a Linear Quadratic Programming (LQP) optimization problem to find the actuation variables (joint torques) which enable to follow some objectives at best (e.g., hand trajectory, balance, postural preferences), while respecting dynamic and biomechanical constraints (e.g., joint position and actuation limits) [17]. More details about the DHM simulation can be found in [9].

D. Action Set Extraction

A naive approach to address the issue of the continuous action space would be to randomly sample a large number n of actions, resulting in a discrete, but very large, action set \mathcal{A}_n . However, while the FH algorithm returns a decision

³We actually consider positive and negative torques separately to represent agonist and antagonist muscles, hence two fatigue values are associated to each joint.

in a bounded time, this time depends on the depth of the horizon h but also on the size of the action set. Planning within \mathcal{A}_n then becomes computationally too expensive for long-term planning (i.e. more than a few steps ahead). Thus, a small-size relevant subset of actions within \mathcal{A}_n needs to be identified.

Since we consider offline action-set extraction, the extracted subset should be relevant for any possible state of the system. But the state space here is continuous (fatigue state), so it is not possible to test for all possible states. We therefore propose to use a fast short-term planning (greedy, FH_1) step within \mathcal{A}_n , to explore the state-space and identify relevant actions for the visited states. The short-term planning is conducted over many task cycles (n_{cycles}), and repeated for a set \mathcal{B} of various initial beliefs b_0 , to cover the state space as exhaustively as possible. The extracted action set \mathcal{G} (*greedy-used*) is the union of all actions selected during the planning. Algorithm 2 summarizes the proposed method.

Algorithm 2 Greedy-Used

Input \mathcal{A}_n \triangleright base action set
Input \mathcal{B} \triangleright a set of initial beliefs
Input n_{cycles} \triangleright number of task cycles
Initialize $\mathcal{G} \leftarrow \emptyset$
for $b_0 \in \mathcal{B}$ **do**
 $\mathcal{G}_{b_0} \leftarrow \emptyset$
 $b \leftarrow b_0$
 for $s \in b$ **do** \triangleright repeat for all possible states in b
 repeat n_{cycles} **times**
 $a \leftarrow FH(1, b)$
 $\mathcal{G}_{b_0} \leftarrow \mathcal{G}_{b_0} \cup \{a\}$
 $(s', o, r) \leftarrow POMDP(s, a)$
 $b(s'')_{s'' \in S} \propto O(s'', o) \sum_{s \in S} T(s, a, s'').b(s)$
 $s \leftarrow s'$
 end
 $\mathcal{G} \leftarrow \mathcal{G} \cup \mathcal{G}_{b_0}$
return \mathcal{G} \triangleright Greedy-Used set derived from \mathcal{B}

III. EXPERIMENTS

In order to evaluate the performance (decision time, obtained rewards, and robustness to initial belief) of action sets obtained following our approach, we benchmarked it on a *simulated* task scenario similar to [9]. This section first describes the task and the POMDP parameters (Section III-A); then the benchmark tests (Section III-B), and finally the metrics and analyses that were used (Section III-C).

A. Simulated Scenario

1) *Task*: The repetitive co-manipulation task depicted in Fig.1 is considered as a test-bed, with $n_{cycles} = 100$ repetitions. Each cycle lasts 16s: the robot brings the object at the given pose, the human reaches to it with its right hand carrying a 5kg manual tool and operates on the object for 6s while maintaining a fixed whole-body posture (no

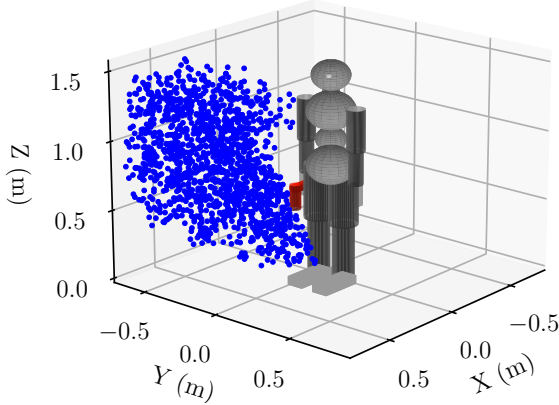


Fig. 2. 3D positions, with respect to the human model, of the actions (i.e. poses of the object brought by the robot) forming the \mathcal{A}_{1500} set. \mathcal{A}_{1500} is a fine discretization of the continuous action space. The tested sets \mathcal{G}_i and \mathcal{R}_k are all extracted from \mathcal{A}_{1500} .

significant human-robot or -object force exchange), then the human returns to its initial posture (idle with arms down), and the next cycle starts.

2) *Reward Function*: The reward function of the POMDP is defined by $r(s') = r(f_{t+1}) = -C(f_{t+1})$ where f_{t+1} is the fatigue state at s' . The fatigue cost $C(f_t)$ is:

$$C(f_t) = \sum_{j \in \text{joints}} \frac{(f_t^j)^2}{2N} \quad (2)$$

with $2N$ the combined number of agonist and antagonist joints (N each) in the human model.

3) *Human Postural Reaction*: Because of its kinematic redundancy, the human can adopt various postural reactions to reach one given object pose. In this example, we use two possible manually defined postural reactions: *back reaction* and *arm reaction*. In *back reaction* the human aims at keeping its upper-arm vertical to reduce shoulder torque, and prefers to use its back when reaching. Conversely, in *arm reaction* the human aims at keeping its back upright to reduce lumbar torque, and prefers to use its arm when reaching. These postural reactions are implemented by tuning parameters in the LQP controller animating the human model [9]. Here we consider that at each cycle, the human chooses the reaction which minimizes its fatigue cost at the end of the cycle. The human reaction is therefore deterministic.

B. Benchmark

In order to test the proposed action set extraction method, we first create the large action set \mathcal{A}_n with $n = 1500$, by randomly sampling from the continuous shared robot and human Cartesian workspace. Only positions are considered, the object orientation is kept constant in this work. The shared workspace is estimated based on the human body dimensions, and the feasibility of each action (3D object pose) is verified through the DHM simulation. Figure 2 shows the sampled positions in \mathcal{A}_{1500} around the digital human.

1) *Influence of the set of initial beliefs*: The greedy-used action set \mathcal{G} is extracted from \mathcal{A}_{1500} by running the FH_1 algorithm over 100 cycles, and for a set \mathcal{B} containing n_b randomly selected initial beliefs. n_b initial beliefs are used to create \mathcal{G} so that the belief space is explored as exhaustively as possible, resulting in a set \mathcal{G} yielding robust performance. Nevertheless, the extracted set \mathcal{G} may depend on the choice of initial beliefs in \mathcal{B} . We therefore repeat the creation of \mathcal{G} using different sets \mathcal{B}_i of initial beliefs, each \mathcal{B}_i giving one greedy-used set \mathcal{G}_i , and we compare the performance of these sets \mathcal{G}_i . Three different types of \mathcal{B}_i are explored:

- $\mathcal{B}_i^{\text{single}}$ is a set of $n_b = 20$ randomly selected single-state initial beliefs (i.e. a single possible initial state, no uncertainty). Twenty randomly selected sets $\mathcal{B}_i^{\text{single}}$ ($0 < i < 20$) are tested, each resulted in one *greedy-used* action set $\mathcal{G}_i^{\text{single}}$.
- $\mathcal{B}_i^{\text{multi}}$ is a set of $n_b = 20$ randomly selected multi-state initial beliefs (i.e. several possible initial states, with uncertainty). Twenty randomly selected sets $\mathcal{B}_i^{\text{multi}}$ ($0 < i < 20$) are tested, each resulted in one *greedy-used* action set $\mathcal{G}_i^{\text{multi}}$. While the initial beliefs are randomly selected, we limit the choice to 3-states initial beliefs (i.e. 3 possible initial states).
- $\mathcal{B}^{\text{fatigue}}$ is a set containing one single-state initial belief for each joint⁴ j , corresponding to the state where joint j is fully fatigued ($f_j = 1$) and the other joints are fully relaxed ($f_{k \neq j} = 0$). $\mathcal{B}^{\text{fatigue}}$ is expected to result in a *greedy-used* action set $\mathcal{G}^{\text{fatigue}}$ that contains at least one action acting to minimize fatigue for each joint.

2) *Comparison with other extraction methods*: We compare the performance of the *greedy-used* action set with 3 other types of action sets. In order to further validate the robustness of the *greedy-used* sets \mathcal{G}_i , we conduct the comparison using all the sets \mathcal{G}_i described in Section III-B.1, except $\mathcal{G}^{\text{fatigue}}$ (i.e. we use 40 sets \mathcal{G}_i : $\mathcal{G}_i^{\text{single}}$ and $\mathcal{G}_i^{\text{multi}}$ for $0 < i < 20$). The other types of action sets are:

- \mathcal{A}_{1500} : The large initial action set. We hypothesize that planning in \mathcal{A}_{1500} results in fatigue mitigation performance better or equal than \mathcal{G} , since there are more actions to choose from. However, the size of \mathcal{A}_{1500} makes it impossible to plan within the time limit of a task cycle with a horizon $h > 1$.
- \mathcal{R}_k : A set of randomly sampled actions (within \mathcal{A}_{1500}) of similar size k as \mathcal{G} . Since there are 40 sets \mathcal{G}_i , the number of actions k in \mathcal{R}_k is chosen as the average number of actions of all sets \mathcal{G}_i . This comparison serves to verify that the fatigue mitigation performance is not merely related to the number of actions in the set. We hypothesize that planning in \mathcal{R}_k results in fatigue mitigation performance lower than \mathcal{G} since the actions constituting \mathcal{G} are selected optimally vs. randomly for \mathcal{R}_k . The decision time will be similar since both sets contain the same number of actions. The performance of \mathcal{R}_k is likely unstable (if $k \ll 1500$), thus 20 randomly

⁴Except joints in the left arm that are omitted since the left arm is not actively used in the considered scenario, hence its fatigue will remain low.

sample 20 sets \mathcal{R}_k^j are included in the benchmark.

- \mathcal{M}_2 : The set containing the two manually selected actions used in [9]. These two actions (low and high poses) were selected to induce significantly different human postures. We hypothesize that planning in \mathcal{M}_2 results in fatigue mitigation performance lower than \mathcal{G} since the \mathcal{M}_2 action set was not optimized.

3) *Evaluation*: The action sets tested, both for the influence of the initial belief set and for the extraction method, are all evaluated in the same way. They are used for planning with the finite horizon algorithm FH_h for different horizons h , on the simulated scenario described in Section III-A. The performance of the different sets are compared in terms of decision time, and fatigue mitigation performance.

The fatigue mitigation performance however depends on the initial state and belief. For each action set and each planning horizon h , the planning and execution of the robot policy is therefore tested for 50 different initial beliefs b_0 . Among them, 25 are single-state initial beliefs, and 25 are multi-state (3-states) initial beliefs. All are randomly sampled. Single-state b_0 serve for testing the approach when the robot has a full knowledge of the human fatigue, whereas multi-state b_0 serve to test the robustness of the performance in the presence of uncertainty. Finally, for all evaluations, the real initial fatigue state is contained within the initial belief b_0 .

C. Data Analysis

1) *Performance Criteria*: Performance is evaluated with respect to both decision time during online planning, and fatigue mitigation. Fatigue mitigation performance is measured with the average fatigue cost over the whole task:

$$\bar{C} = -\frac{1}{n_{cycles}} \sum_{k=1}^{n_{cycles}} r_k. \quad (3)$$

2) *Statistical Tests*: The fatigue mitigation performance is compared for the different action sets using statistical tests. The performance is analyzed separately when testing from single-state vs. multi-state initial beliefs (25 data points for each action set in both cases). For each type of action set, we use the fatigue mitigation performance obtained with the highest planning horizon h , which decision time still remains under the task cycle time⁵. Data are tested for normality using the Shapiro-Wilk test. Some conditions do not follow the normality assumption, hence non-parametric Wilcoxon signed-rank tests are used to perform pairwise comparisons.

Each *condition* (i.e. type of action set extraction method, or type of initial belief set used to run greedy-used) contains several repetitions. We therefore conduct pairwise comparisons between all pairs of action sets obtained with a different method (e.g., \mathcal{A}_{1500} vs. all \mathcal{G}_i^{single} and all \mathcal{G}_i^{multi} ; all \mathcal{R}_k vs. all \mathcal{G}_i^{single} and all \mathcal{G}_i^{multi} ; \mathcal{M}_2 vs. all \mathcal{G}_i^{single} and all

\mathcal{G}_i^{multi}). In the results, we report the percentage of pairwise comparisons for which the p-value is $p < 0.0001$. Statistical analyses are conducted in Python.

IV. RESULTS

A. Size of the greedy-used action sets

The *greedy-used* action set extraction approach resulted in action sets \mathcal{G} containing 23.6 ± 2.1 actions when \mathcal{G} was derived from single-state initial beliefs (\mathcal{G}_i^{single}), 19.5 ± 1.6 actions when \mathcal{G} was derived from multi-state initial beliefs (\mathcal{G}_i^{multi}), and 32 actions for $\mathcal{G}^{fatigue}$ (derived from 64 initial beliefs). Compared to the initial set \mathcal{A}_{1500} containing 1500 actions, the greedy-used extraction approach allows to divide the size of the action set by 65 (\mathcal{G}_i^{single}) to 75 (\mathcal{G}_i^{multi}) in average.

B. Influence of the set of initial beliefs

Fig. 3 shows one representative example of the fatigue mitigation vs. online decision time performance graph for the different initial belief sets \mathcal{B}_i used for action set extraction, and for different planning horizons. The performance distribution within \mathcal{G}_i^{single} (resp. within \mathcal{G}_i^{multi}) is narrow. So, overall, the performance within one type of initial belief set \mathcal{B}_i is relatively stable, suggesting that $n_b = 20$ randomly selected initial beliefs in \mathcal{B}_i are sufficient to yield a robust action set. Then, the sets \mathcal{G}_i^{single} derived from single-state sets \mathcal{B}_i^{single} and \mathcal{G}_i^{multi} derived from multi-state sets \mathcal{B}_i^{multi} have close performance, both in terms of decision time and fatigue mitigation. When the planning horizon increases, the decision time is slightly larger for \mathcal{G}_i^{single} sets. This is in agreement with the larger average number of actions in \mathcal{G}_i^{single} (23.6 ± 2.1) compared to \mathcal{G}_i^{multi} (19.5 ± 1.6). Conversely, the performance of $\mathcal{G}^{fatigue}$ extracted from a hand-crafted set of initial beliefs is lower, both in terms of decision time (since this set contains more actions) and of fatigue mitigation.

Table I summarizes the results of the statistical comparisons between each pair of action sets extracted from different type of initial belief sets. For each evaluated action set, the performance used for the statistical comparison is the one obtained with the longest planning horizon, which decision time remains below the task cycle time. When evaluating on multi-state initial beliefs, all \mathcal{G}_i^{single} and \mathcal{G}_i^{multi} sets outperform the $\mathcal{G}^{fatigue}$ set. \mathcal{G}_i^{single} and \mathcal{G}_i^{multi} sets are not significantly different from each other. When evaluating on single-state initial beliefs, all \mathcal{G}_i^{multi} sets and about half \mathcal{G}_i^{single} sets outperform the $\mathcal{G}^{fatigue}$ set. \mathcal{G}_i^{multi} sets outperform the \mathcal{G}_i^{single} sets in about half the comparisons. Thus it seems that action sets extracted using a set \mathcal{B}_i of multi-state initial beliefs yield the best and most robust performance. However, the observed difference in fatigue mitigation performance may also partly be due to the planning horizon. Indeed, the planning horizon is one step shorter for $\mathcal{G}^{fatigue}$ and for some \mathcal{G}_i^{single} sets with single-state initial belief evaluations.

⁵We actually used a time limit of 12s, which is a few seconds shorter than the full cycle time, to allow the robot to observe the postural reaction of the human when the human starts moving from its neutral posture at the beginning of the cycle.

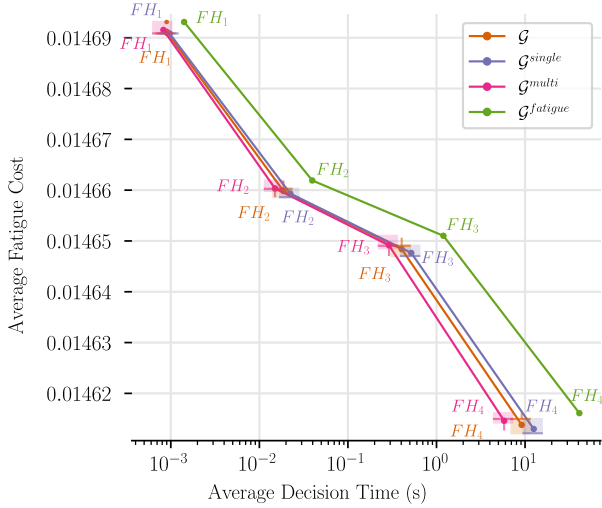


Fig. 3. Fatigue mitigation performance vs. online decision time, for the action sets extracted from the different sets \mathcal{B}_i of initial beliefs. \mathcal{G}_i^{single} (20 sets) corresponds to single-state initial belief sets \mathcal{B}_i^{single} , \mathcal{G}_i^{multi} (20 sets) to multi-state initial belief sets \mathcal{B}_i^{multi} , and $\mathcal{G}_i^{fatigue}$ (1 set) to a hand crafted set of initial beliefs $\mathcal{B}_i^{fatigue}$. For \mathcal{G}_i^{single} and \mathcal{G}_i^{multi} , the distribution across the 20 sets is shown in blue (resp. magenta), and the indicated decision time is the average over all 20 sets. Colored dots correspond to average values, and horizontal lines to medians. \mathcal{G} (orange line) corresponds to the distribution of performance across all sets \mathcal{G}_i^{single} and \mathcal{G}_i^{multi} together. The performance of each action set was evaluated with different planning horizons FH_h . Data shown on the graph correspond to one representative evaluation, i.e. when testing from one initial belief out of the 25 single-state initial beliefs used for the benchmark.

C. Comparison with other extraction methods

The results of the previous section show that there is little difference between the performance of \mathcal{G}_i^{single} and \mathcal{G}_i^{multi} sets. Hence, in this section these sets are all analyzed together and denoted as \mathcal{G}_i (40 sets). The randomly selected sets \mathcal{R}_k were then designed to contain 22 actions, which was the average size of the 40 action sets \mathcal{G}_i .

Fig. 4 shows one representative example of the fatigue mitigation vs. online decision time performance graph for the different methods used to extract the action set (\mathcal{A}_{1500} , \mathcal{G}_i , \mathcal{R}_{22} , and \mathcal{M}_2), and for different planning horizons. Sets extracted with *greedy-used* (\mathcal{G}_i) and with random selection (\mathcal{R}_{22}) have similar decision time, since they contain approximately the same number of actions. However, the fatigue mitigation performance is better with \mathcal{G}_i sets, showing that it is not merely the number of actions in the set that matters. The \mathcal{M}_2 set has a low decision time since it contains only 2 actions. But it has the worst performance regarding fatigue mitigation, even with long planning horizons. Hence, while the planning horizon plays a role in the fatigue mitigation performance (it is beneficial to plan with a longer horizon), it needs to be associated to a relevant and rich-enough action set. The full set \mathcal{A}_{1500} and the *greedy-used* sets \mathcal{G}_i have close –and best overall– performance in terms of fatigue mitigation. However, for a same planning horizon, the decision time of \mathcal{A}_{1500} is much longer given the large

TABLE I

PERCENTAGE OF COMPARISONS WHERE THE AVERAGE FATIGUE COST OF ACTION SETS IN ROWS IS SIGNIFICANTLY LOWER THAN SETS IN COLUMNS. THE THRESHOLD FOR SIGNIFICANCE IS 0.0001. NUMBERS IN PARENTHESES INDICATE THE PLANNING HORIZON USED TO MEASURE THE PERFORMANCE.

	Evaluation with single-state b_0		
	\mathcal{G}_i^{single} ($h=3.5 \pm 0.5$)	\mathcal{G}_i^{multi} ($h=4 \pm 0$)	$\mathcal{G}_i^{fatigue}$ ($h=3 \pm 0$)
\mathcal{G}_i^{single}	–	0.0	55.0
\mathcal{G}_i^{multi}	45.0	–	100.0
$\mathcal{G}_i^{fatigue}$	0.0	0.0	–
	Evaluation with multi-state b_0		
	\mathcal{G}_i^{single} ($h=3 \pm 0$)	\mathcal{G}_i^{multi} ($h=3 \pm 0$)	$\mathcal{G}_i^{fatigue}$ ($h=2 \pm 0$)
\mathcal{G}_i^{single}	–	0.0	100.0
\mathcal{G}_i^{multi}	0.0	–	100.0
$\mathcal{G}_i^{fatigue}$	0.0	0.0	–

number of actions it contains.

Table II summarizes the results of the statistical comparisons between each pair of action sets extracted with a different method. For each evaluated action set, the performance used for the statistical comparison is the one obtained with the longest planning horizon such that its decision time remains below the task cycle time. All *greedy-used* sets \mathcal{G}_i significantly outperform all other sets, whether the evaluation is conducted with single-state or multi-state initial beliefs. Interestingly, \mathcal{G}_i outperforms \mathcal{A}_{1500} with respect to fatigue mitigation, even though all actions in \mathcal{G}_i are also in \mathcal{A}_{1500} . This is likely due to the small size of \mathcal{G}_i , which allows to plan with a longer horizon, whereas \mathcal{A}_{1500} is limited to a horizon of 1 to remain compatible with the task cycle time. This shows that increasing the planning horizon is beneficial with respect to fatigue mitigation. Overall, these results show that the *greedy-used* action set extraction method gives sets that are a good compromise between the number of actions on one hand, and their diversity and relevance with respect to fatigue mitigation on another hand.

TABLE II

PERCENTAGE OF COMPARISONS WHERE THE AVERAGE FATIGUE COST OF ACTION SETS IN ROWS IS SIGNIFICANTLY LOWER THAN SETS IN COLUMNS. THE THRESHOLD FOR SIGNIFICANCE IS 0.0001. NUMBERS IN PARENTHESES INDICATE THE PLANNING HORIZON USED TO MEASURE THE PERFORMANCE.

	Evaluation with single-state b_0			
	\mathcal{A}_{1500} ($h=1 \pm 0$)	\mathcal{G} ($h=3.8 \pm 0.4$)	\mathcal{M}_2 ($h=11 \pm 0$)	\mathcal{R}_{22} ($h=4 \pm 0$)
\mathcal{A}_{1500}	–	0.0	100.0	100.0
\mathcal{G}	100.0	–	100.0	100.0
\mathcal{M}_2	0.0	0.0	–	0.0
\mathcal{R}_{22}	0.0	0.0	100.0	–
	Evaluation with multi-state b_0			
	\mathcal{A}_{1500} ($h=1 \pm 0$)	\mathcal{G} ($h=3 \pm 0$)	\mathcal{M}_2 ($h=7 \pm 0$)	\mathcal{R}_{22} ($h=2.9 \pm 0.3$)
\mathcal{A}_{1500}	–	0.0	100.0	100.0
\mathcal{G}	100.0	–	100.0	100.0
\mathcal{M}_2	0.0	0.0	–	0.0
\mathcal{R}_{22}	0.0	0.0	100.0	–

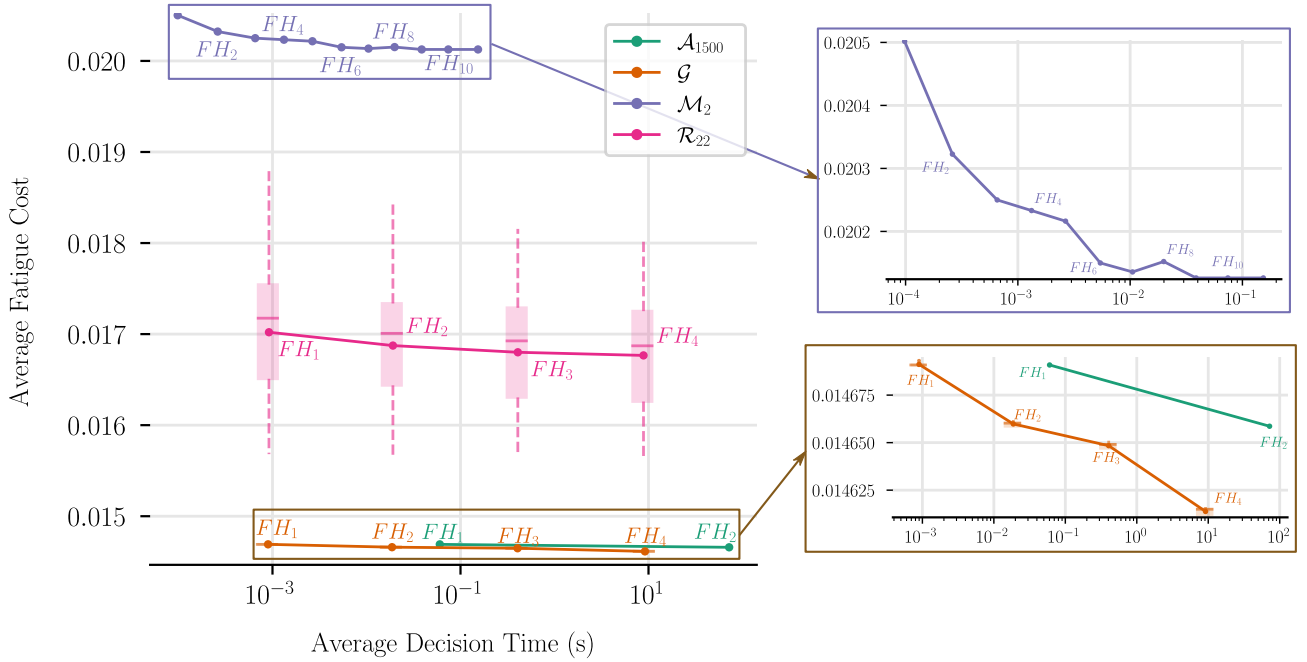


Fig. 4. Fatigue mitigation performance vs. online decision time, for the action sets extracted with the different methods: full set (\mathcal{A}_{1500} , 1 set), *greedy-used* extraction (\mathcal{G}_i , 40 sets), random selection (\mathcal{R}_{22} , 20 sets), and manually crafted set (\mathcal{M}_2 , 1 set). For \mathcal{G}_i (resp. \mathcal{R}_{22}), the distribution across the 40 (resp. 20) sets is shown in orange (resp. magenta), and the indicated decision time is the average over all sets. Colored dots correspond to average values, and horizontal lines to medians. The performance of each action set was evaluated with different planning horizons FH_h . Data shown on the graph correspond to one representative evaluation, i.e. when testing from one initial belief out of the 25 single-state initial beliefs used for the benchmark.

V. DISCUSSION

The results of this study suggest that the proposed *greedy-used* approach is relevant to extract a small subset of relevant actions from a large action space, in the context of fatigue mitigation with a cobot. The *greedy-use* extraction approach allowed to divide the size of the action set by about 70, which enabled to use a longer horizon for online planning compared to the large initial set representing the continuous space. This is especially important, as increasing the planning horizon has been shown beneficial to improve fatigue mitigation performance, both here and in previous work [9]. Importantly, when considering a same planning horizon, the average fatigue cost of *greedy-used* action sets was close to that of the initial large set. Thus, reducing the size of the action set was not detrimental to the fatigue mitigation performance, which confirms the relevance of the extracted actions.

While reducing the action set size is crucial for online planning applications, a diversity of actions to choose from should nevertheless be maintained. Indeed, sets of randomly selected actions (\mathcal{R}_{22}) performed better than the set of manually selected actions (\mathcal{M}_2), even though this latter action set was constituted with fatigue mitigation in mind. The limited diversity offered by only two action choices (compared to the 22 actions in \mathcal{R}_{22}) was therefore not sufficient to allow for an efficient variation of physical load distribution on the different human joints over time. The proposed *greedy-used* approach therefore offers a good trade-off between size of the action set, and diversity of actions in the set.

One advantage of the *greedy-used* approach is that the number of parameters to tune is limited, and their meaning is understandable. The important point is that the belief space be explored as exhaustively as possible during the extraction phase, in order to select suitable responses (i.e. actions) to any possible situation that may happen. This is why repeating the extraction process for several initial beliefs was proposed. The results actually suggest that the extracted action set is more robust (better performance on a diversity of evaluations despite less actions) when using several randomly selected initial beliefs (\mathcal{G}_i^{single} and \mathcal{G}_i^{multi}), compared to using manually defined initial beliefs designed to represent a wide variety of fatigue states ($\mathcal{G}^{fatigue}$). In addition, sets derived from multi-state initial beliefs (\mathcal{G}_i^{multi}) marginally outperformed those derived from single-state initial beliefs (\mathcal{G}_i^{single}). This could be explained by the fact that multi-state beliefs intrinsically consider uncertainty on the fatigue state, hence actions that generate a safer behavior are selected (i.e. actions that result from a trade-off between risk and performance). These safer actions may be more robust on fatigue states that were not explored in the action set extraction phase, compared to actions tailored for one specific deterministic state⁶.

Despite the promising results of the *greedy-used* approach, this work has some limitations. We tested the performance of the *greedy-used* approach on one single co-manipulation

⁶When starting from a single-state initial belief, the fatigue state remains deterministic throughout the whole task, since in the present example the human postural reaction is deterministic, hence the whole evolution is.

scenario, and we cannot exclude that the results do not generalize to other scenarios. For instance, the *greedy-used* extraction approach may result in a set that contains too many actions to be compatible with time constraint of online planning for repetitive co-manipulation tasks. If so, the *greedy-used* approach would need to be extended to further reduce the number of actions in the set. This may require to rank the actions, and therefore assign a value to each action. The number of times an action is selected could be one solution, however the magnitude of its consequences should also be considered. In addition, we did not evaluate the performance of the action sets in case of an erroneous belief, i.e. a belief that does not contain the real fatigue state of the human. Such case would benefit from actions that – along with performance– optimize the information provided about the human state, in order to quickly refine the belief [18].

Another limitation concerns the reliability of the human model, which affects the estimated fatigue evolution and hence performance. The human postural reactions (motion preferences) were manually designed. We chose them to be sensible w.r.t. how a human could react, but we cannot guarantee that they match real human behaviors. In addition, the simulated human always chose the postural reaction resulting in the smallest fatigue cost at the end of the current cycle. In reality, humans may apply some hysteresis in their choices: while the fatigue level definitely affects the choice, humans may also tend to keep the same reaction as before up to a certain level of fatigue. In the future, we therefore plan to identify human motion preferences from real human data. We will also evaluate the performance of the extracted action sets in a real human-robot experiment replicating the simulated scenario. This will be an opportunity to assess whether the human model used for the action set extraction is good enough for the performance to transfer from simulation to real.

VI. CONCLUSION

In this work, we addressed the question of action set extraction from a continuous space, in the context of human fatigue mitigation in highly repetitive human-robot co-manipulation tasks. Long-term planning of robot actions is beneficial to optimize human fatigue mitigation. However, the constraint on the decision time imposed by short task cycle requires that the set of actions used for planning be of limited size. Yet, actions constituting the set should allow for a diversity of fatigue distributions on the different human joints, and be robust to uncertainties on the human state. We proposed a method for offline action set extraction, based on short-term planning and physics-based digital human simulation. The short-term planning phase is conducted to explore the fatigue belief space and simultaneously identify, from a large action space, optimal actions with respect to fatigue mitigation for each visited belief. The digital human simulation serves to evaluate the fatigue induced by each possible action of the robot. We evaluated our *greedy-used* approach in a simulated co-manipulation scenario, and

showed that the action sets thereby extracted are a good compromise between the number of actions in the set, and their diversity and relevance with respect to fatigue mitigation. Importantly, the small size of the resulting set allowed to increase the online planning horizon, which was beneficial for fatigue mitigation. In the future, we will improve the realism of the human model, leveraging data-driven approaches. We also intend to validate our approach on a real human subject experiment, and test it on different human-robot co-manipulation scenarios.

REFERENCES

- [1] J. De Kok, P. Vroonhof, J. Snijders, G. Roullis, M. Clarke, K. Peereboom, P. van Dorst, and I. Isusi, “Work-related musculoskeletal disorders: prevalence, costs and demographics in the eu,” *European agency for safety and health at work*, vol. 1, 2019.
- [2] K. Summers, K. Jinnett, and S. Bevan, “Musculoskeletal disorders, workforce health and productivity in the united states,” *The center for worked health and performance*. London: Lancaster university, 2015.
- [3] L. Punnett and D. H. Wegman, “Work-related musculoskeletal disorders: the epidemiologic evidence and the debate,” *Journal of electromyography and kinesiology*, vol. 14, no. 1, pp. 13–23, 2004.
- [4] D. Srinivasan and S. E. Mathiassen, “Motor variability in occupational health and performance,” *Clinical biomechanics*, vol. 27, no. 10, pp. 979–993, 2012.
- [5] W. Kim, J. Lee, L. Peternel, N. Tsagarakis, and A. Ajoudani, “Anticipatory robot assistance for the prevention of human static joint overloading in human–robot collaboration,” *IEEE Robot. Autom. Lett.*, vol. 3, no. 1, pp. 68–75, 2017.
- [6] B. Busch, G. Maeda, Y. Mollard, M. Demangeat, and M. Lopes, “Postural optimization for an ergonomic human-robot interaction,” in *2017 IEEE/RISJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 2778–2785.
- [7] L. Peternel, C. Fang, N. Tsagarakis, and A. Ajoudani, “A selective muscle fatigue management approach to ergonomic human-robot co-manipulation,” *Robotics and Computer-Integrated Manufacturing*, vol. 58, pp. 69–79, 2019.
- [8] M. Zolotas, R. Luo, S. Bazzi, D. Saha, K. Mabulu, K. Kloeckl, and T. Padir, “Imposing motion variability for ergonomic human-robot collaboration,” *IIEE Transactions on Occupational Ergonomics and Human Factors*, vol. 12, no. 1-2, pp. 123–134, 2024.
- [9] A. Yaacoub, V. Thomas, F. Colas, and P. Maurice, “A probabilistic model for cobot decision making to mitigate human fatigue in repetitive co-manipulation tasks,” *IEEE Robot. Autom. Lett.*, 2023.
- [10] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Prentice Hall, 2010.
- [11] A. Couëtoux, J.-B. Hoock, N. Sokolovska, O. Teytaud, and N. Bonnard, “Continuous upper confidence trees,” in *Learning and Intelligent Optimization*. Springer Berlin Heidelberg, 2011, pp. 433–445.
- [12] M. Hoerger, H. Kurniawati, D. Kroese, and N. Ye, “Adaptive discretization using voronoi trees for continuous-action pomdps,” 2022.
- [13] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, “Planning and acting in partially observable stochastic domains,” *Artificial intelligence*, vol. 101, no. 1-2, pp. 99–134, 1998.
- [14] D. Silver and J. Veness, “Monte-carlo planning in large pomdps,” *Advances in neural information processing systems*, vol. 23, 2010.
- [15] National Research Council and Institute of Medicine, *Musculoskeletal Disorders and the Workplace: Low Back and Upper Extremities*. Washington, DC: The National Academies Press, 2001.
- [16] L. Ma, W. Zhang, D. Chablat, F. Bennis, and F. Guillaume, “Multi-objective optimisation method for posture prediction and analysis with consideration of fatigue effect and its application case,” *Computers & Industrial Engineering*, vol. 57, no. 4, pp. 1235–1246, 2009.
- [17] J. Salini, V. Padois, and P. Bidaud, “Synthesis of complex humanoid whole-body behavior: A focus on sequencing and tasks transitions,” in *2011 IEEE international conference on robotics and automation*. IEEE, 2011, pp. 1283–1290.
- [18] M. Araya, O. Buffet, V. Thomas, and F. Charpillet, “A pomdp extension with belief-dependent rewards,” in *Advances in Neural Information Processing Systems*, vol. 23. Curran Associates, Inc., 2010.