Probing the Functional Role of Muscle Synergies in Reinforcement Learning-Based Torso Balance via Neural Reconstruction and Synergy Ablation

1st Siyuan Liu, 2nd Jiahao Chen*
1. State Key Laboratory of Multimodal Artificial Intelligence Systems,
Institute of Automation, Chinese Academy of Sciences, Beijing, China
2. School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China
{liusiyuan2023, jiahao.chen}@ia.ac.cn

***Abstract*—Muscle synergies are widely hypothesized as an organizational principle for controlling highly redundant musculoskeletal systems. However, most existing synergy analyses rely on post-hoc statistical decompositions of recorded muscle activity, making it difficult to directly relate synergy function to closed-loop control. In this work, we propose a reinforcement learning-based neuromuscular control framework that embeds muscle synergies as explicit, policy-structured components within the control architecture. A synergy decoder maps low-dimensional modulation signals to high-dimensional muscle excitations, enabling synergy structure to emerge as an intrinsic part of the learned controller. Based on this, we introduce a decoder-based synergy analysis method using neural reconstruction and synergy ablation. We evaluate the proposed approach on a high-dimensional musculoskeletal torso balance task. Synergy function is probed through both open-loop replay ablation and closed-loop online ablation. The results show that reconstruction-based importance and functional importance are related but not equivalent. It suggests a division of labor among synergy pathways, with different synergy channels contributing to feasibility, terminal stabilization, and precision shaping. Overall, this work provides a control-relevant framework for studying the functional roles of muscle synergies and bridges biomechanical motor control with reinforcement learning. The experiment video can be found in: https://sites.google.com/view/syn-decoder/.**

***Index Terms*—musculoskeletal simulation, reinforcement learning, torso balance, neural ablation, muscle synergy**

I. Introduction

As a central component of the human musculoskeletal system, the torso plays a crucial role in whole-body movement and postural stability [1]. Compared with the highly dexterous limbs, whose broad range of motion and diverse functions have attracted extensive attention, the torso is often perceived as functionally simpler and has therefore been relatively less studied. Nevertheless, the torso is actuated by hundreds of small, densely arranged muscles with complex anatomical organization, and it provides essential mechanical support for balance and stabilization [2]. A deeper understanding of torso motor control can thus offer valuable insights into how complex embodied musculoskeletal systems achieve robust and coordinated behavior [3], [4].

One influential hypothesis for explaining control in such highly redundant muscle systems is the concept of muscle synergies [5]. This concept posits that the central nervous system does not control individual muscles independently but instead recruits muscles in coordinated groups. Under this view, muscle activation patterns can be approximated as combinations of time-invariant synergy weights and time-varying synergy activation coefficients [6]. However, most existing synergy analyses rely on post-hoc statistical decompositions of recorded muscle activity, such as autoencoder [7], Nonnegative Matrix Factorization (NMF) [8] and Principal Component Analysis (PCA) [9]. This makes it difficult to directly relate the extracted synergy structure to the underlying control pipeline and to causally probe the functional roles of individual synergy pathways. Accordingly, there is a need for methods that explicitly connect synergy representations with closed-loop motor control.

NMF and PCA provide plausible explanations for coordinated muscle recruitment by revealing group-wise structure in muscle activity. However, statistical explained variance does not necessarily correspond to functional importance within the control pipeline. Even when an apparently compact or optimal synergy representation can be identified, systematically testing the functional role of muscle synergies under closed-loop control remains largely unexplored. Probing how functional synergies contribute to muscle recruitment within a neuromuscular control loop is therefore essential for advancing our understanding of how complex muscle systems are regulated during movement.

Recent advances in high-fidelity musculoskeletal simulation [10], [11] and reinforcement learning (RL) have opened new avenues for studying neuromuscular control from a control-centric perspective [12], [13]. Within this paradigm, neuromuscular controllers can be learned through interaction with physics-based musculoskeletal models, enabling the synthesis of closed-loop control policies that operate directly in high-dimensional muscle activation spaces [14]. However, most

*Corresponding author: *Jiahao Chen.*
This work was supported by the National Natural Science Foundation of China under Grant 62573410.

prior studies have focused primarily on the resulting muscle behaviors produced by learned controllers, while paying relatively little attention to the internal structure of the neuromuscular control policies themselves [15]. As a result, post-hoc analyses of biomechanical or electrophysiological data remain the dominant approach for studying muscle coordination, and how muscle synergies functionally contribute to control within a closed-loop neuromuscular controller remains largely unexplored.

To address this gap, we embed muscle synergies as structured components within the neural network control architecture. This provides a new pathway for jointly investigating motor control strategies and muscle coordination. Rather than inferring synergies post hoc from recorded activation data, we integrate a synergy decoder directly into the policy. This method allows synergy structure to emerge as an intrinsic part of the learned controller itself and makes it directly amenable to functional analysis. The proposed framework enables direct access to synergy recruitment within neuromuscular control and supports multi-faceted evaluation of synergy function through neural reconstruction and targeted synergy ablation. Together, this approach establishes a principled link between biomechanical motor control and reinforcement learning, offering new insights into the organization and functional roles of muscle synergies in complex musculoskeletal systems.

II. Methodology

A. Musculoskeletal Torso Model and Balance Task

In this work, we introduce the torso musculoskeletal model [16] developed in MyoSuite to train a torso balance control agent as shown in Fig. 1. The model comprises 210 muscle–tendon units (MTUs) actuating 18 independent degrees of freedom (DoFs): flexion–extension, lateral bending, and axial rotation at each lumbar joint. It features a rigid pelvis and sacrum, five articulated lumbar vertebrae, and a rigid upper torso representing a lumped thoracic spine and rib cage, making it well suited for studying the principles of human torso motion and balance control. Furthermore, in this work we employ a dedicated balance environment built on top of this model, in which the observation state at time t is defined as $\boldsymbol{s}_t = \{\boldsymbol{q}_t, \dot{\boldsymbol{q}}_t, \hat{\boldsymbol{q}}, \boldsymbol{e}_t\}$. This state consists of the joint positions $\boldsymbol{q}_t$, joint velocities $\dot{\boldsymbol{q}}_t$, the desired balance target configuration

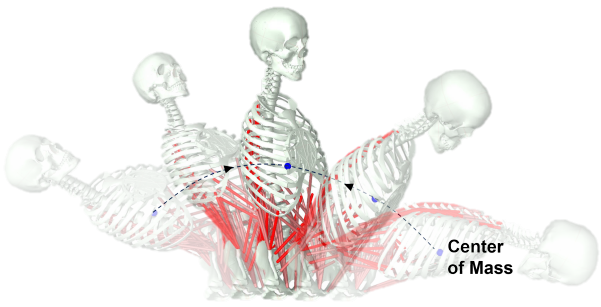

Fig. 1. Musculoskeletal torso model and torso balance task setup.

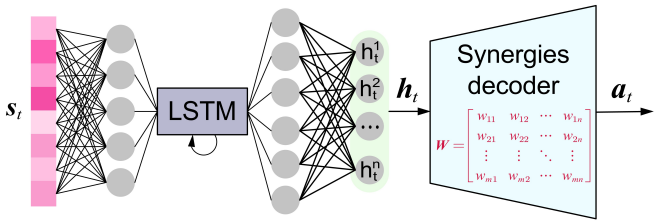

Fig. 2. LSTM-based neural network controller with explicit muscle synergies decoder, where the muscle modulation coefficients $\boldsymbol{h}_t$ are mapped to muscle excitations $\boldsymbol{a}_t$ through the muscle synergy weight matrix $\boldsymbol{W}$.

$\hat{\boldsymbol{q}}$ and the balance pose error $\boldsymbol{e}_t = \hat{\boldsymbol{q}} - \boldsymbol{q}_t$. Correspondingly, the environment action is represented by a 210-dimensional vector of muscle excitations, each component bounded between -1 and 1. The environment provides a reward that encourages minimizing the Euclidean error between the current and target joint angles, as defined by the following equation:

$$r_t = \exp\left(-5\|\boldsymbol{e}_t\|_2\right) \tag{1}$$

At the beginning of each episode, the initial state of the musculoskeletal model is sampled from the range of joint configurations. The control objective is to drive the torso model to the balanced posture shown in the middle of Fig. 1. Each episode lasts for 200 timesteps, corresponding to 2 seconds.

To quantitatively evaluate balance performance, we analyze the trajectory of the model's center of mass (CoM) over each episode and its deviation from the desired CoM target. Specifically, if the mean error e_f between the CoM trajectory and the target CoM over the final 0.5s of an episode falls below a threshold φ, the corresponding torso balance trial is considered successful.

B. Synergy-Structured Neural RL Agent

To model muscle synergy structure within the neuromuscular controller, as shown in Fig. 2, we employ a single linear synergy decoder that maps synergy modulation coefficients $\boldsymbol{h}_t$ to muscle excitations $\boldsymbol{a}_t$, which are computed as follows:

$$\boldsymbol{a}_t = \boldsymbol{W}\boldsymbol{h}_t \tag{2}$$

where $\boldsymbol{W} \in \mathbb{R}^{m\times n}$ denotes the muscle synergy matrix, and m and n denotes the number of muscle and synergies, respectively.

In this neuromuscular controller, the output of the LSTM-based policy network just before the synergy decoder is interpreted as the muscle modulation coefficients $\boldsymbol{h}_t$. The synergy decoder thus realizes a network-based synergy module that is formally equivalent to conventional muscle synergy theory [8], [17], with the column vector $\boldsymbol{W}_{:,j}$ representing the j-th muscle synergy in the synergy weight matrix.

Furthermore, to achieve more sample-efficient reinforcement learning in the high-dimensional muscle action space, we employ a latent exploration method [18] to enhance the exploration capability of vanilla Proximal Policy Optimization

(PPO) [19]. Consequently, the muscle action during training is defined as follows:

$$\widetilde{\boldsymbol{a}}_t = \boldsymbol{a}_t + \mathcal{N}(0, \boldsymbol{W}\Sigma_h\boldsymbol{W}^\top + \Sigma_a) \tag{3}$$

where Σ_h is a positive-definite diagonal matrix that perturbs the synergy modulation state, and Σ_a is the original independent action noise covariance used in PPO.

Thus, the muscle excitation $\boldsymbol{a}_t$ follows a multivariate Gaussian distribution as a function of the synergies decoder weight and the covariance matrix as follows:

$$\begin{aligned}\log \pi_\theta(\widetilde{\boldsymbol{a}}_t|\boldsymbol{s}_t) = &-\frac{N_a}{2}\log(2\pi) - \frac{1}{2}\log|\boldsymbol{W}\Sigma_h\boldsymbol{W}^\top + \Sigma_a| \\ &- \frac{1}{2}(\widetilde{\boldsymbol{a}}_t - \boldsymbol{W}\mathbf{x})^\top(\boldsymbol{W}\Sigma_h\boldsymbol{W}^\top + \Sigma_a)(\widetilde{\boldsymbol{a}}_t - \boldsymbol{W}\mathbf{x})\end{aligned} \tag{4}$$

Based on this, the probability of each action depends explicitly on the synergy decoder weight matrix. This explicit dependence allows the learned muscle synergy structure to actively shape exploration during reinforcement learning and facilitates the discovery of efficient muscle recruitment strategies in the redundant muscle action space.

C. Decoder-based Synergy Analysis Method

Here we propose a decoder-based muscle synergy analysis approach. Unlike conventional NMF or PCA, which extract synergies through a purely linear decomposition, the synergy weights in our neural decoder naturally form primitive spatiotemporal patterns of multi-muscle activity. Instead of refitting a new synergy matrix on recorded muscle activation data, we directly reuse the synergy weights learned by the motor controller and analyze their contribution in the context of the trained policy. For the trained neural network controller as mentioned in Sec. II-B, both the synergy modulation amplitudes and the resulting muscle excitations applied to the musculoskeletal model can be directly recorded. The neuromuscular control process can therefore be described for all time steps $t = 0, \ldots, T-1$ as follows:

$$\boldsymbol{A} = \boldsymbol{W}\boldsymbol{H} \tag{5}$$

where $\boldsymbol{A} \in \mathbb{R}^{m\times T}$ denotes the muscle excitations of m muscles over T time steps, and $\boldsymbol{H} \in \mathbb{R}^{n\times T}$ denotes the synergy modulation amplitudes of n muscle synergies over the same T time steps.

Furthermore, to quantify the reconstruction contribution associated with individual muscle synergies in the decoder, the R^2-value for each synergy is defined using a synergy ablation method as follows:

$$R_j^2 = 1 - \frac{\sum_{t=0}^{T-1}||\boldsymbol{a}_t - \hat{\boldsymbol{a}}_t^{(j)}||_2}{\sum_{t=0}^{T-1}||\boldsymbol{a}_t - \bar{\boldsymbol{a}}||_2} \tag{6}$$

where $\bar{\boldsymbol{a}}$ denotes the mean muscle excitation over the trial, and $\hat{\boldsymbol{a}}_t^{(j)}$ denotes the reconstructed muscle excitation at time t, obtained by feeding the ablated synergy modulation vector $\mathring{\boldsymbol{h}}_t^{(j)} = [0, \ldots, 0, h_t^{(j)}, 0, \ldots, 0]^\top$, in which only the j-th synergy coefficient is retained and all others are set to zero, into the synergy decoder. This corresponds to single-synergy reconstruction, where the neuromuscular controller is forced to operate through a single synergy channel. Thus, we obtain a neural network–based approach to quantify the reconstruction contribution of each synergy channel within the learned controller, following a rationale analogous to classical lesion studies in neuroscience, where disrupting specific brain regions is used to assess their contribution to behavior. In practice, we rank synergies by their individual R_j^2 values and construct cumulative reconstruction curves to assess how many high-ranked synergies are sufficient to account for most of the observed muscle excitations.

III. Experiment

A. Torso Balance Performance

As shown in Fig. 3(a), we trained the synergy-structured RL agent. Compared with the standard PPO baseline, the synergy-structured RL approach converges faster and attains a slightly higher final return. As an important metric for investigating torso balance, the model's CoM is projected onto the ground plane to track subtle deviations from the target equilibrium. The dynamic movements of the musculoskeletal torso during 200 balance episodes are demonstrated in Fig 3(b). The CoM trajectories generally start from the torso workspace and converge to the target balance range following a curved path. It is notable that some trajectories, particularly those observed in the left region of the image, represent unsuccessful balance trials. This failure may be attributed to the musculoskeletal system reaching the limits of its joint range of motion, where the available torque generated by the torso muscle is insufficient to restore balance without the assistance of limb movements.

Specifically, to evaluate the torso balance performance in different directions, a specific balance experiment was conducted with initial states equally spaced on a circle with the same radius R. The 0° direction represents the initial state in which the model's CoM is displaced by a distance R anteriorly (i.e., along the positive x-axis from the balanced CoM). The 90° direction represents the initial state in which the model's CoM is displaced by a distance R to the left (i.e., along the positive y-axis) from the balanced CoM. As shown in Fig. 3(c), the CoM trajectories initialized at evenly spaced locations on the same circle all converge to the desired Balanced CoM position. The mean final reaching error e_f is 4.795 ± 0.002 mm, and all eight trials are deemed successful according to the success metric defined in II-A. This indicates that the proposed neural controller effectively exploits the learned synergies weights and modulates them online to accomplish the torso-balancing task from different initial directions.

B. Functional Analysis of Neural Synergy Structure

Furthermore, the muscle synergy aptitudes and muscle excitations from the balancing experiments with eight evenly spaced initial directions were recorded for functional analysis of different synergies structures. Following the procedure in Sec. II-C, the R_j^2 of each individual synergy was computed and

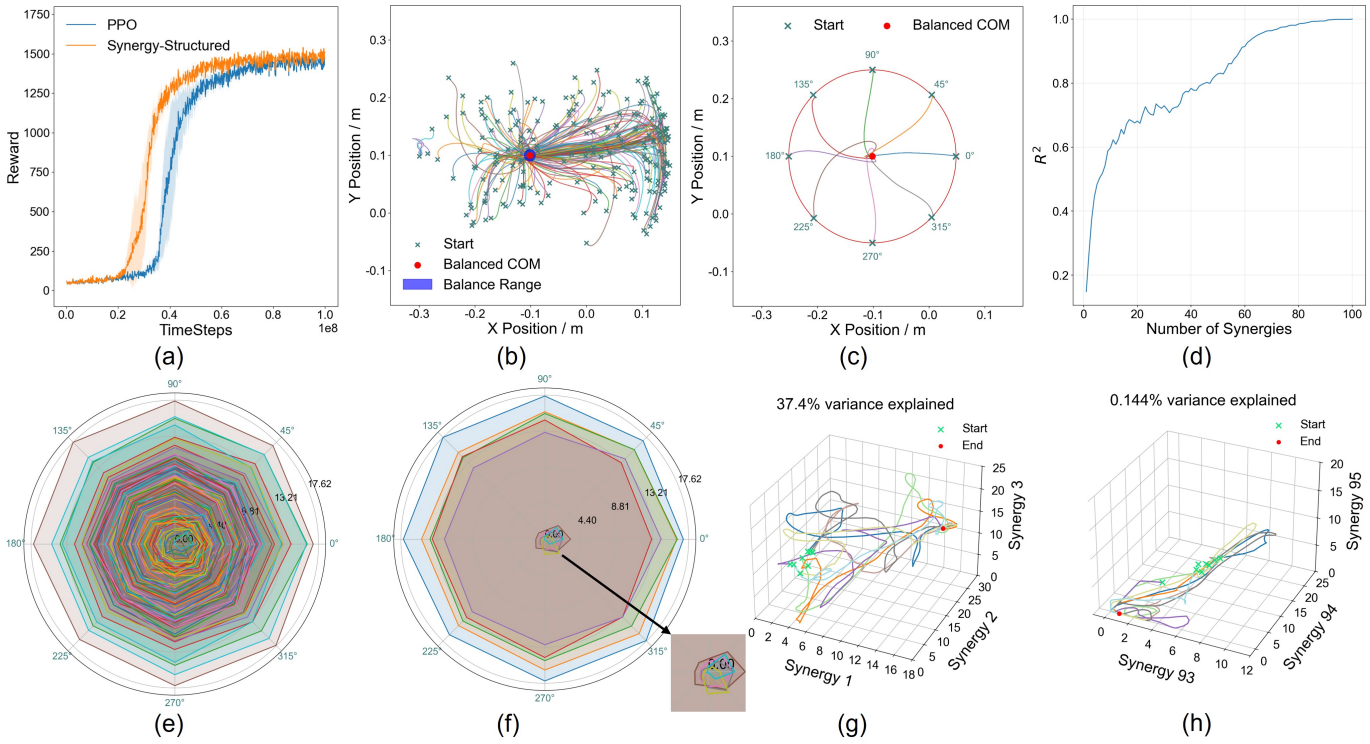

Fig. 3. (a). Learning Curves of the RL Agents. Comparison of the average reward achieved by the LSTM-based Policy with the explicit Synergy Decoder across different numbers of muscle synergies n. The shaded area represents the variability of the policies across 3 different random seeds. (b). CoM Trajectories during 200 torso balance episodes. (c). CoM trajectories for balance trials initiated from 8 evenly spaced initial CoM positions on a circle of equal radius. The angle labels represent the initial direction relative to the balanced posture COM. (d). Cumulative Muscle Excitation Reconstruction Performance. The R^2 value for reconstructing the original muscle excitations using an increasing number of ranked muscle synergies. Synergies are ranked from the most important to the least important based on their individual contribution R_j^2. (e). Single Synergy Modulation Amplitudes across initial Directions. Radar plot showing the activation amplitude of each individual synergy channel across balance trials initiated from the 8 different starting directions. The radial axes represent the average activation aptitude for each trial. (f). Polar plot comparing the average activation amplitude of the five most functionally important synergies and the five least functionally important synergies across the 8 initial directions. (g). Synergy recruitment trajectories in the subspace of most 3 important synergies. (h). Synergy recruitment trajectories in the subspace of least 3 important synergies.

used for ranking. Then, according to the resulting importance order of R_j^2, the cumulative muscle excitation reconstruction was obtained by progressively restoring the contribution of synergies one by one. As shown in Fig. 3(d), the overall R^2 curve exhibits an increasing trend as more synergies are included in the reconstruction, and it approaches nearly 100% explained variance when the number of synergies reaches 95. This suggests that each synergy channel in the synergies decoder plays a specific role in generating the muscle excitations.

It is worth noting that the R^2 curve is not strictly monotonic and is substantially lower than that obtained by PCA-based analysis. This is because our synergy analysis is designed to evaluate, under the currently learned control policy, the relative importance of each synergy channel in generating the observed muscle excitation patterns, rather than to provide an optimal linear explanation of muscle excitations in the classical sense (e.g., PCA or NMF). In particular, the learned synergies in the controller are not strictly orthogonal. Nevertheless, this neural reconstruction-based analysis offers a deeper perspective for interpreting the underlying neuromuscular control mechanisms.

To analyze the recruitment of different synergies during the balancing task under different directions, we use the average synergy modulation during the single direction balancing process, which is defined as follows:

$$\bar{H}_i^d = \frac{\sum_{t=0}^{T-1} h_{it}^d}{T} \tag{7}$$

where h_{it}^d denotes the activation coefficient of the i-th synergy at time step t during the balancing process in direction d, and $\bar{H}_i^d$ represents the average synergy modulation of the i-th synergy for balance in direction d.

Thus, the average synergy modulation for balancing tasks under different directions is visualized using radar plots. As shown in Fig. 3(e), the majority of synergies exhibit weak directional selectivity, as they are recruited to a similar extent across movements from different directions. A more specific illustration for the five most important and the five least important synergies is provided in Fig. 3(f). For the first five synergies, strong recruitment is observed for balancing motions in almost all directions. In contrast, the last five synergies show pronounced directional specificity: they are selectively recruited only during balancing motions initiated from the anterior side of the body, while motions starting from

posterior directions rely minimally on these synergies. This finding is consistent with previous synergy analyses, which reported that synergies with lower explained variance tend to exhibit stronger task-specific or directional selectivity [20], [21].

C. Synergy Ablation Analysis

To further probe the functional roles of individual synergy channels during torso balance, we conducted two synergy ablation experiments. The first experiment was performed on the eight-direction balance task introduced in Sec. III-A. For each direction, we recorded the synergy modulation amplitudes $\boldsymbol{H}$ produced by the intact controller. During the ablation test, we set the selected synergy channel(s) in the recorded modulation sequence to zero and fed the modified modulation amplitudes into the synergy decoder. The resulting decoder output was then applied to the torso model as an open-loop replay command, thereby isolating the behavioral consequences of removing the contribution of the selected synergy pathway without allowing online feedback compensation.

Based on the R^2-ranking defined in Sec. III-B, we ablated the highest-ranked and the lowest-ranked synergy channels, respectively. As shown in Fig. 4(a), ablating the top-ranked synergy led to failure across all initial directions. Notably, the CoM trajectories initially followed patterns similar to those of the intact controller, but diverged abruptly near the final stage when approaching the target balanced CoM, resulting in a late-stage loss of stabilization around the balanced CoM. This observation suggests that the top-ranked synergy plays a critical role in maintaining equilibrium in the vicinity of the target state. In contrast, ablating the lowest-ranked synergy (Fig. 4(b)) also produced noticeable degradation in the terminal behavior, indicating that even a synergy channel with low reconstruction-based importance can be functionally relevant for achieving accurate end-state stabilization under open-loop replay.

The second ablation experiment was conducted in a closed-loop manner. During the eight-direction torso balance task, the selected synergy channels were clamped to zero online at each control step, thereby ablating the corresponding pathways while still allowing the policy to react to the evolving state. This setup enables us to evaluate the behavioral contribution of specific synergy channels under feedback control, where potential compensatory mechanisms across channels may emerge.

In this experiment, we first ablated the top five synergy channels according to the R^2-based ranking as described in Sec. III-B. Unlike the open-loop replay ablation, closed-loop ablation of the top five synergies did not lead to outright task failure. As shown in Fig. 4(c), the terminal CoM positions exhibited only a small offset from the balanced CoM. The mean final reaching error increased to 11.776 ± 0.045 mm, indicating degraded steady-state precision. This result suggests that synergy channels exhibit functional complementarity: under closed-loop feedback, the controller can partially compensate for the suppressed pathways by redistributing modulation across the remaining synergies, preserving task feasibility at the expense of reduced accuracy.

We further evaluated closed-loop ablation of the bottom five synergies. As shown in Fig. 4(d), the CoM trajectories were affected to a smaller extent, and the mean final reaching error increased only slightly to 4.923 ± 0.014 mm. Compared to ablating the top five synergies, ablating the bottom five produced a markedly smaller degradation in precision, suggesting that these low-ranked synergy channels are more replaceable by other channels in the decoder, although they may still contribute to fine-grained aspects of the control policy. Although the functions of individual synergy channels can be partially substituted by other (redundant) channels in the decoder under closed-loop feedback, it remains meaningful to investigate the unique contributions of the top-ranked synergies. As shown in Fig. 5, we compared the intact controller against the controller with the top five synergies ablated. While both controllers eventually achieved torso balance, ablating the top five synergies led to substantially larger oscillations during the error-correction phase and a higher steady-state CoM error. These results indicate that under closed-loop feedback, task feasibility can be maintained even after ablating top-ranked synergies. Nevertheless, these synergies are essential for well-damped error dynamics and high-precision steady-state balance. In other words, compensation by the remaining synergies is possible, yet it comes at the cost of increased oscillation and reduced final accuracy. This suggests that top-ranked synergies disproportionately contribute to damping and precision, whereas remaining synergies can compensate only partially.

As shown in Fig. 6, progressive closed-loop ablation further characterizes how functional sensitivity is distributed across synergy channels. We sequentially clamped synergy channels to zero while keeping the controller in closed loop. Two ablation orders were compared: a forward order that removes synergies from high to low reconstruction-based rank, and a reverse order that removes them from low to high rank. The average balance error remains small when ablating low-ranked synergies first, but increases much earlier when ablating high-ranked synergies first. In our experiments, noticeable degradation emerges after removing roughly the first 20 synergies in the forward order, whereas the reverse order requires removing roughly 40 synergies to reach a comparable increase in error. This asymmetric robustness curve indicates that redundancy in the synergy space is not uniform: a relatively small subset of top-ranked synergies carries a disproportionate share of the stabilizing and precision-critical control structure, while many low-ranked synergies are more replaceable under feedback.

Overall, reconstruction-based importance and closed-loop functional importance are related but not equivalent. The open-loop replay ablation and the closed-loop online ablation jointly reveal the dynamical roles of synergy pathways in torso balance control. Specifically, open-loop ablation exposes terminal-stage failures when compensation is unavailable, whereas closed-loop ablation reveals partial compensation accompanied by degraded control quality. These findings suggest

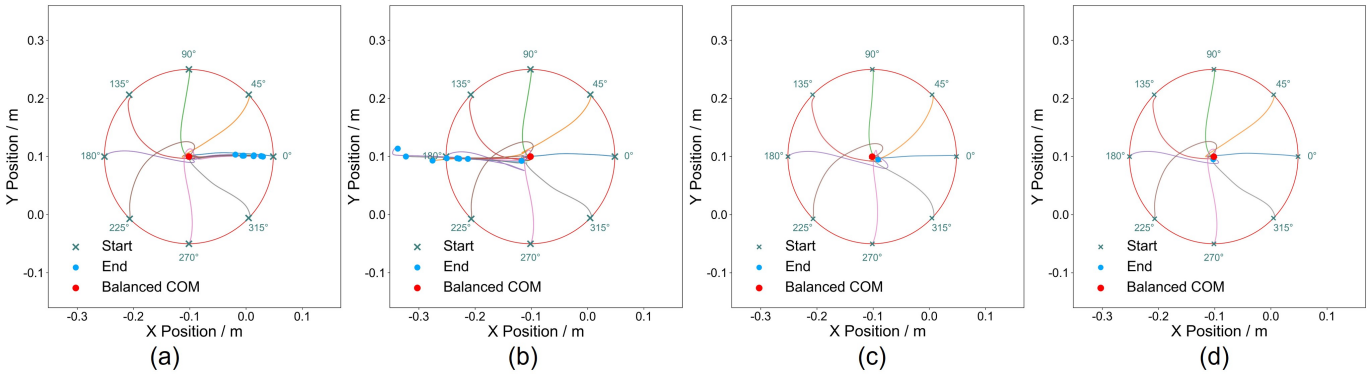

Fig. 4. (a). CoM trajectories under open-loop replay with the top-ranked synergy channel ablated. (b). CoM trajectories under open-loop replay with the bottom-ranked synergy channel ablated. (c). CoM trajectories under closed-loop online ablation with the top five synergy channels clamped to zero at every control step. (d). CoM trajectories under closed-loop online ablation with the bottom five synergy channels clamped to zero at every control step.

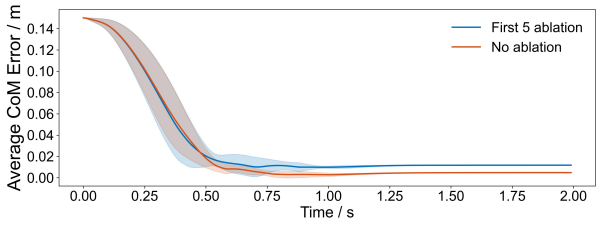

Fig. 5. Torso balance CoM error correction dynamics comparing the intact controller and the controller with the top five synergies ablated. Solid curves show the mean CoM error averaged across balance trials initiated from the eight evenly spaced directions, and the shaded regions indicate the standard deviation across directions.

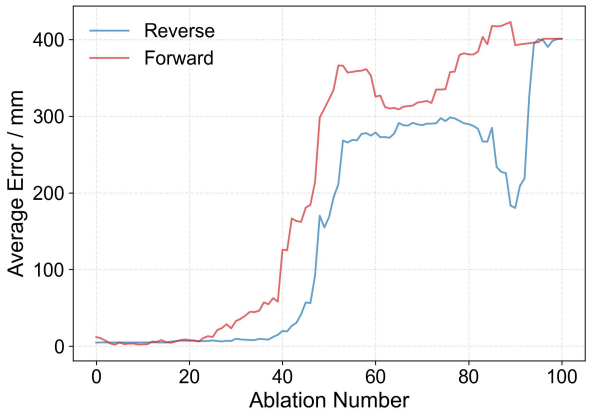

Fig. 6. Closed-loop progressive synergy ablation. Synergy channels are sequentially clamped to zero during closed-loop control. Forward ablation removes synergies from high to low reconstruction-based rank, whereas Reverse ablation removes synergies from low to high rank. Curves show the average balance error as a function of the number of ablated channels.

that synergy pathways support balance through a division of labor, where different channels contribute differently to coarse recovery, terminal stabilization, and precision shaping of the closed-loop response.

IV. Discussion

Unlike post-hoc decomposition methods such as PCA and NMF, which are widely used for muscle synergy analysis, the decoder synergies in our framework are policy-embedded primitives that directly parameterize both action generation and exploration through the structured controller. This decoder-based neuromuscular controller therefore provides a complementary perspective on musculoskeletal motor control: it enables the direct study of control-relevant synergies that emerge within the trained policy, rather than purely statistical synergies extracted from recorded activation patterns.

Despite this methodological shift, our findings exhibit notable parallels with prior observations from statistical synergy analyses. In the open-loop replay ablation, we found that synergy channels with low reconstruction importance can still produce measurable behavioral consequences, suggesting that low explained-variance components are not necessarily functionally irrelevant. This is consistent with previous reports that low-variance components can be task-related and reflect meaningful structure in motor behavior [20], [21]. More broadly, our ablation results highlight that reconstruction-based importance and behavioral importance are not equivalent: synergy channels that contribute little to statistical reconstruction can still be important for achieving successful control.

Moreover, the closed-loop ablation experiments reveal another fundamental property of biological motor systems: motor redundancy. Our results indicate that redundancy also exists in the synergy space: under feedback control, the policy can partially compensate for missing channels by redistributing modulation across remaining synergies. This observation is conceptually related to analyses of uncontrolled manifolds that characterize families of solutions compatible with task success [22]. At the same time, the redundancy is limited: ablating high-ranked synergies leads to degraded stabilization quality, implying that certain synergy pathways make disproportionate contributions to precision and stability. Progressive closed-loop ablation further shows that this redundancy is not uniformly distributed: performance degrades substantially

earlier when high-ranked synergies are removed first than when low-ranked synergies are removed first, indicating a concentrated stabilizing core. Together, these results support a division-of-labor view in which different synergy pathways contribute differently to feasibility, terminal stabilization, and precision shaping.

This work also has several limitations that motivate future research. First, the capacity of the synergy decoder is chosen manually, and the impact of decoder dimensionality on learned synergy structure and functional specialization warrants systematic investigation. Second, the current torso balance task is an isolated stabilization setting that lacks external interactions and disturbances; incorporating perturbations or interaction-rich scenarios would provide a more stringent test of synergy function under ecological constraints. Finally, the proposed analysis should be validated across other motor paradigms (e.g., reaching, locomotion, and whole-body tasks) to assess its generality beyond torso balance.

V. Conclusion

This work investigated the functional roles of muscle synergies in a reinforcement learning–based neuromuscular controller for a high-dimensional torso balance task. By embedding muscle synergies as explicit, policy-structured components through a synergy decoder, we enabled direct access to synergy modulation within the control process. The results show that the proposed framework yields synergy structures consistent with observations from traditional synergy analyses, while revealing that statistical importance and functional importance are related but not equivalent. The synergy pathways with low statistical importance can still exert significant functional effects in the torso balance task. Furthermore, closed-loop ablation experiments demonstrate partial compensation across synergy channels, indicating redundancy and complementarity within the synergy space. However, ablating top-ranked synergies leads to increased transient oscillations and reduced steady-state precision, suggesting that these channels play a critical role in stabilizing the closed-loop response. Together, these findings support a division-of-labor perspective, in which different synergy pathways contribute differently to feasibility, terminal stabilization, and precision shaping of balance control. Overall, the proposed approach provides a control-relevant framework for probing muscle synergies and strengthens the connection between biomechanical motor control and reinforcement learning.

References

[1] N. Kavcic, S. Grenier, and S. M. McGill, "Determining the stabilizing role of individual torso muscles during rehabilitation exercises," *Spine*, vol. 29, no. 11, pp. 1254–1265, 2004.

[2] A. Moya-Esteban, H. van der Kooij, and M. Sartori, "Robust estimation of lumbar joint forces in symmetric and asymmetric lifting tasks via large-scale electromyography-driven musculoskeletal models," *Journal of biomechanics*, vol. 144, p. 111307, 2022.

[3] C. Zuo, K. He, J. Shao, and Y. Sui, "Self model for embodied intelligence: Modeling full-body human musculoskeletal system and locomotion control with hierarchical low-dimensional representation," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 13 062–13 069.

[4] P. Schumacher, D. Haeufle, D. Büchler, S. Schmitt, and G. Martius, "Dep-rl: Embodied exploration for reinforcement learning in overactuated and musculoskeletal systems," in *The Eleventh International Conference on Learning Representations*.

[5] Y. Chen, Y. Wang, F. Gao, B. Dong, W. Chen, and R. Liu, "A computational method to identify optimal functional muscle synergies from estimated muscle activations," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–14, 2024.

[6] D. M. Mulla and P. J. Keir, "Neuromuscular control: from a biomechanist's perspective," *Frontiers in sports and active living*, vol. 5, p. 1217009, 2023.

[7] D. Buongiorno, C. Camardella, G. D. Cascarano, L. Pelaez Murciego, M. Barsotti, I. De Feudis, A. Frisoli, and V. Bevilacqua, "An undercomplete autoencoder to extract muscle synergies for motor intention detection," in *2019 International Joint Conference on Neural Networks (IJCNN)*, 2019, pp. 1–8.

[8] M. F. Rabbi, C. Pizzolato, D. G. Lloyd, C. P. Carty, and L. E. Diamond, "Non-negative matrix factorisation is the most appropriate method for extraction of muscle synergies in walking and running," *Scientific Reports*, vol. 10, no. 1, 2020.

[9] K. Zhao, Z. Zhang, H. Wen, B. Liu, J. Li, A. d'Avella, and A. Scano, "Muscle synergies for evaluating upper limb in clinical applications: A systematic review," *Heliyon*, vol. 9, no. 5, 2023.

[10] H. Wang, V. Caggiano, G. Durandau, M. Sartori, and V. Kumar, "Myosim: Fast and physiologically realistic mujoco models for musculoskeletal and exoskeletal studies," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 8104–8111.

[11] S. Liu, B. Jiang, L. Han, S. Zhong, C. Song, H. Liu, and J. Chen, "Biosyngrasp: Bio-inspired structured reinforcement learning with synergetic exploration and human demonstration for dexterous grasping of musculoskeletal robot," *IEEE Transactions on Automation Science and Engineering*, vol. 23, pp. 6102–6116, 2026.

[12] M. Rv and S. Rakshit, "Deep reinforcement learning based control of lower limb exoskeleton," in *2024 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2024, pp. 1–6.

[13] H. Yu, Z. Huang, Q. Liu, I. Carlucho, and M. S. Erden, "Replication of impedance identification experiments on a reinforcement-learning-controlled digital twin of human elbows," in *2024 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2024, pp. 1–8.

[14] L. Han, L. Cheng, M. Ma, F. Yang, and H. Li, "Muscle synergy-guided reinforcement learning for embodied musculoskeletal motion skill learning," *IEEE transactions on bio-medical engineering*, 2026.

[15] H. Wang, J. Kovecses, and G. Durandau, "Reinforcement learning identifies age-related balance strategy shifts," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 33, pp. 4078–4088, 2025.

[16] R. Walia, M. Billot, K. Garzon-Aguirre, S. Subramanian, H. Wang, M. I. Refai, and G. Durandau, "Myoback: A musculoskeletal model of the human back with integrated exoskeleton," in *2025 International Conference On Rehabilitation Robotics (ICORR)*, 2025, pp. 128–135.

[17] X. Guo, S. Huang, B. He, C. Lan, J. J. Xie, K. Y. S. Lau, T. Takei, A. D. P. Mak, R. T. H. Cheung, and K. Seki, "Inhibitory components in muscle synergies factorized by the rectified latent variable model from electromyographic data," *IEEE Journal of Biomedical and Health Informatics*, vol. 29, no. 2, pp. 1049–1061, 2025.

[18] A. S. Chiappa, A. Marin Vargas, A. Huang, and A. Mathis, "Latent exploration for reinforcement learning," *Advances in Neural Information Processing Systems*, vol. 36, pp. 56 508–56 530, 2023.

[19] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[20] Y. Yan, J. M. Goodman, D. D. Moore, S. A. Solla, and S. J. Bensmaia, "Unexpected complexity of everyday manual behaviors," *Nature communications*, vol. 11, no. 1, p. 3564, 2020.

[21] A. S. Chiappa, P. Tano, N. Patel, A. Ingster, A. Pouget, and A. Mathis, "Acquiring musculoskeletal skills with curriculum-based reinforcement learning," *Neuron*, vol. 112, no. 23, pp. 3969–3983, 2024.

[22] T. Bennett, L. Thomas, and A. D. Wilson, "Affordances for throwing: An uncontrolled manifold analysis," *Plos one*, vol. 19, no. 4, p. e0301320, 2024.