LEARNING FUNCTIONAL SPECIALISATION FOR BIMANUAL ROBOTIC CONTROL

Anonymous authors
Paper under double-blind review

ABSTRACT

Bimanual robotic control often requires diverse functional demands to be handled simultaneously under challenging control conditions. Modular controllers provide one way to distribute these demands, but how control responsibilities should be organised across controllers and the two arms, and how this organisation affects performance and flexibility, remains unclear. Biological motor control provides an organisational principle in which each hemisphere predominantly controls the contralateral limb, alongside functional specialisation and stable hand preferences. We therefore investigate whether analogous organisational principles can benefit modular robotic motor control in a musculoskeletal system with two communicating recurrent controllers. Using a 2×2 design, we compare differential (controller-specific) and shared functional objectives together with fixed and balanced (role-varying) hand–task assignments. We find that task-aligned functional specialisation is strongest when differential objectives are combined with fixed hand–task assignment aligned with contralateral control routing. This condition achieves the highest performance and control efficiency on the task, reaching 81.33% bimanual success with significant improvements over the other training conditions. Its performance advantage over shared objectives increases under stronger motor noise but narrows under stronger energy penalties. It also shows lower hand-role flexibility: reversing the hand roles decreases bimanual success by 80.46 percentage points under differential objectives with fixed hand–task assignment, compared with only 1.51 percentage points under shared objectives with balanced hand–task assignment. Our findings show that, in bimanual robotic control with each controller more strongly routed to one hand, functional specialisation can develop through the organisation of learning responsibility and is associated with advantages in performance and control efficiency, whereas more distributed training objectives and task assignments favour greater interchangeability when roles change.

1 INTRODUCTION

Bimanual behaviour often relies on a division of labour between the two hands rather than identical contributions (Swinnen, 2002). In bimanual robotic manipulation, the two arms often take complementary roles, with one arm providing stabilisation while the other performs the primary manipulation (Liu et al., 2025; Yu et al., 2025). In human bimanual control, handedness is associated with stable preferences for using particular hands for particular motor functions (Woytowicz et al., 2018). Such stable hand–function associations provide repeated experience in the same role, reinforcing learned advantages for recurring control demands (Provins, 1997; Marcori et al., 2019). At the neural level, bimanual control is also organised asymmetrically: the two cerebral hemispheres make different task-dependent contributions to motor control, with each hemisphere predominantly controlling the contralateral limb while retaining weaker ipsilateral influence (Serrien et al., 2006; Brinkman & Kuypers, 1973). This organisation establishes a stable hemisphere–limb correspondence while allowing each hemisphere to influence both limbs. The two hemisphere–limb systems also exhibit complementary functional specialisations: the dominant system has been associated with movement dynamics and trajectory control, whereas the non-dominant system shows advantages in positional stabilisation (Sainburg, 2005; Mutha et al., 2012). Such functional specialisation supports a division of labour in which different systems contribute according to their respective

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

strengths, allowing complementary control demands to be handled by different parts of the system instead of requiring equal involvement in every aspect of the task.

Modular neural controllers can support functional differentiation, but modularity alone does not determine how responsibilities should be distributed (Wolpert & Kawato, 1998; Haruno et al., 2001; Michaels et al., 2020). Different training objectives encourage components to develop distinct functional priorities (Dobs et al., 2022; Rinaldo et al., 2025), while repeated practice reinforces hand-specific motor competencies (Marcori et al., 2019; Dexheimer et al., 2024; Arac et al., 2026). These mechanisms have largely been studied separately, leaving unclear how these two forms of organisation interact during learning and when the resulting specialisation improves behaviour.

Theoretical accounts suggest that the benefits of specialised organisation are expressed differently under different control pressures (Seoane, 2023). In motor control, signal-dependent noise causes execution variability to increase with motor-command magnitude (Harris & Wolpert, 1998), while energetic cost creates pressure to achieve task goals with lower motor expenditure (Todorov & Jordan, 2002). Interhemispheric communication is also delayed: transcallosal conduction between motor cortical areas has been estimated at approximately 13 ms (Meyer et al., 1995), and precisely timed interhemispheric interactions contribute to successful bimanual coordination (Liuzzi et al., 2011). These considerations motivate testing how motor noise, energetic cost, and inter-controller communication delay influence the behavioural benefits of learned functional specialisation.

Taken together, these observations motivate us to incorporate these organisational principles into bimanual robotic motor control. We study how controller-level functional objectives and stable hand–task assignment interact to produce functional specialisation, and how the resulting organisation affects performance, efficiency, and hand-role flexibility under control and task conditions. To do so, we use a bilateral recurrent controller in which each controller primarily controls the contralateral hand while retaining weaker influence over the ipsilateral hand. This fixed controller–hand mapping provides a stable foundation for examining how controller-level functional responsibility interacts with hand-role assignment. The controller operates in a musculoskeletal bimanual task combining trajectory tracking and positional stabilisation. Our main contributions are as follows:

- We introduce a controlled 2×2 framework that disentangles controller-level functional objectives, defined as either separate optimisation of the two controllers for trajectory tracking and positional stabilisation or a shared objective combining these two functional demands, from hand–task assignment within a bimanual task where one hand tracks a moving target while the other stabilises against external perturbations.
- We show that specialisation is strongest when controller-specific objectives are combined with fixed hand–task assignment. This organisation is accompanied by higher bimanual performance and efficiency, with the benefit of controller-specific objectives increasing under stronger motor noise and decreasing under stronger energy penalties. For robotic motor control, these results suggest that functional responsibility across modular controllers should be considered together with effector roles rather than designed independently.
- We show that stronger functional specialisation can be associated with reduced hand-role flexibility when functional demands are reassigned across hands. Controllers trained with shared functional objectives and varied hand–task assignments perform better when hand roles change, revealing different balances between role-specific performance and flexibility across changing bimanual task configurations.

2 Related work

Modularity and specialisation in robot learning. Computational models have examined functional and hemispheric specialisation across cognitive domains, including language and visual processing (Chang & Lambon Ralph, 2020; Behrmann et al., 2026; Dobs et al., 2022). In robot learning, modular architectures have been used to distribute task or functional demands across specialised components, including separate motor-control policies, task-specific modules, and expert-based manipulation policies (Hu et al., 2023; Jian et al., 2023; Li et al., 2025; Deng et al., 2026). Functional differentiation has also been studied in bilateral motor control, where controllers trained with different functional objectives can develop complementary preferences for reaching and positional stabilisation (Rinaldo et al., 2025). Related work has examined how communication contributes to

performance in bilateral modular controllers under control constraints (Li et al., 2026). These studies do not establish whether controller-specific objectives alone are sufficient, or how they interact with stable hand–task experience to produce specialisation and affect motor-control performance, efficiency, and flexibility.

Learning responsibility across multiple objectives. Multi-task learning can improve generalisation by sharing information across related tasks, but optimising multiple tasks together can also produce negative transfer when task objectives compete (Zhang & Yang, 2022; Navon et al., 2022; Ban & Ji, 2024). Existing approaches address this tension through task-specific components and explicit task routing (Ding et al., 2023), specialised expert allocation (Chen et al., 2023; Shazeer et al., 2017; Fedus et al., 2022), or optimisation methods that balance competing task losses (Liu et al., 2023). These approaches primarily determine how learning signals and model capacity are distributed across tasks or model components, without separately considering whether the same functional demands are repeatedly associated with particular effectors during training. It therefore remains unclear whether learned functional specialisation arises from controller-specific objectives, repeated hand–function experience, or their interaction. Our work separates these two sources of learning responsibility within the same bilateral control architecture.

3 Learning Setup and Measurements

3.1 Simulated Musculoskeletal Bimanual System and Tasks

We constructed a simulated bimanual control system in MotorNet (Codol et al., 2024) using two RigidTendonArm26 musculoskeletal effectors (Fig. 1A). The two arms were controlled simultaneously and were not mechanically coupled. The bimanual task comprised two simultaneous task components: holding and tracking. In the holding task, one hand maintained a fixed target while resisting external force perturbations with magnitudes sampled from 0.5–1.5 N. In the tracking task, the other hand followed a continuously moving target with frequency sampled from 0.75–1.25 Hz (Fig. 1A). Further environment and task details are provided in Appendix A.1.

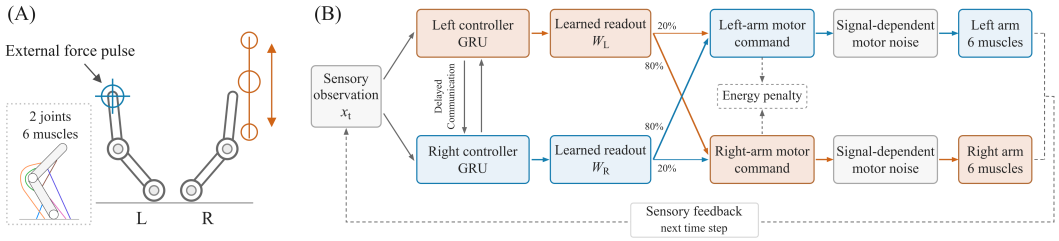

Figure 1: **(A)** Bimanual task and musculoskeletal effectors. One hand stabilises a fixed target against external force perturbations (left) while the other tracks a continuously moving target (right). The inset shows the MotorNet RigidTendonArm26 effector used for each arm, a planar two-degree-of-freedom model actuated by six Hill-type muscles. Two identical effectors are arranged symmetrically in the workspace. **(B)** Bilateral modular controller architecture. Two GRU-based controllers receive shared sensory observations, exchange information through delayed bidirectional connections, and contribute motor commands to the two arms. Motor commands are subject to signal-dependent noise, with motor-command magnitude determining the energetic penalty during training.

3.2 Bilateral Modular Controller Architecture

Inspired by human motor control, in which the two hemispheres interact during bimanual coordination (Serrien et al., 2006; Pauwels & Gooijers, 2023), we designed a bilateral controller comprising two recurrent modules representing interacting hemispheric controllers (Fig. 1B). Each module was implemented as a gated recurrent unit (GRU) (Cho et al., 2014) with 64 hidden units and received the full 46-dimensional sensory observation of the bimanual system. The controllers maintained separate recurrent states and exchanged learned signals bidirectionally with communication delay d.

Each controller also projected its hidden state through a learned motor readout to produce commands for all twelve muscle actuators. To reflect the predominance of contralateral motor control together

with weaker ipsilateral contributions (Brinkman & Kuypers, 1973), the two controller outputs were combined using fixed 80:20 contralateral-to-ipsilateral routing at the motor-readout stage, before the controller contributions were summed and transformed into muscle commands. This routing structure was held fixed across all training conditions and therefore acted as a common architectural prior, allowing both controllers to contribute to both arms while assigning stronger gains to contralateral outputs. Further architecture and control-constraint details are provided in Appendix A.2.

3.3 Learning Responsibility Design

We varied learning responsibility across two dimensions (Fig. 2): **(A) training objectives**, which determine which functional losses update each controller, and **(B) task assignments**, which determine the hand–function mappings presented during training. The forward control architecture remained unchanged across all conditions.

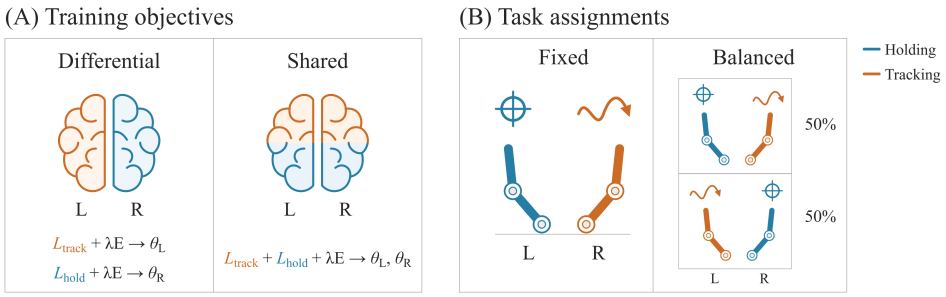

Figure 2: Experimental design for learning responsibility. **(A)** Training-objective structure. Differential training routes tracking and holding losses separately to the two controllers, whereas Shared training applies the combined functional loss to both controllers, with the energy term common to both schemes. **(B)** Task-assignment structure. Fixed training maintains the same hand–task mapping throughout training, whereas balanced training presents right-tracking/left-holding and left-tracking/right-holding mappings equally across trials.

Training objectives. Let $\mathcal{L}_{\text{track}}$, $\mathcal{L}_{\text{hold}}$, and $\mathcal{L}_{\text{energy}}$ denote the tracking, holding, and energetic losses. Both objective schemes included the same energetic loss, weighted by λ. Under Differential training, the tracking and holding objectives were assigned to the left and right controller parameter groups, respectively, with gradients restricted to the corresponding group (Fig. 2A):

$$\mathcal{J}_L = \mathcal{L}_{\text{track}} + \lambda \mathcal{L}_{\text{energy}}, \tag{1}$$

$$\mathcal{J}_R = \mathcal{L}_{\text{hold}} + \lambda \mathcal{L}_{\text{energy}}. \tag{2}$$

Under Shared training, both task losses were combined in a single objective whose gradients updated all trainable parameters:

$$\mathcal{J}_{\text{shared}} = \mathcal{L}_{\text{track}} + \mathcal{L}_{\text{hold}} + \lambda \mathcal{L}_{\text{energy}}. \tag{3}$$

Task assignment. Fixed training maintained the right-tracking/left-holding canonical mapping, whereas Balanced training presented the canonical and reversed left-tracking/right-holding mappings equally across trials (Fig. 2B).

Crossing these factors produced four conditions: Differential-Fixed, Differential-Balanced, Shared-Fixed, and Shared-Balanced. Further objective and gradient-assignment details are in Appendix A.3.

3.4 Control Constraints

To evaluate these learning-responsibility schemes under control constraints, we tested three biologically motivated factors: **(1) Energy penalty weight** λ controlled the contribution of the motor-output penalty during training. Because generating motor commands incurs energetic cost, λ determines how strongly energy use is weighted relative to task performance (Todorov & Jordan, 2002). We used $\lambda = 0$ as a no-energy-penalty baseline and logarithmically spaced non-zero values $\{10^{-5}, 10^{-4}, 10^{-3}\}$ for progressively stronger relative penalties. **(2) Motor-noise gain** k controlled

signal-dependent perturbations applied to motor commands, with perturbation amplitude proportional to the motor command and scaled by k (Harris & Wolpert, 1998; Hamilton & Wolpert, 2002). Higher k therefore increases execution noise and reduces command-to-output reliability. This tests unreliable motor execution, relevant to degraded motor function and robotic component degradation. We included $k = 0$ as a deterministic baseline and evaluated stronger noise levels at $k \in \{0.05, 0.1, 0.2\}$. **(3) Communication delay** d determined the delay applied to bidirectional inter-controller signals, motivated by transcallosal conduction delays and the temporal requirements of interhemispheric coordination (Meyer et al., 1995; Liuzzi et al., 2011), and was evaluated at $d \in \{0, 5, 10, 20\}$ simulation steps (50, 100, and 200 ms). These factors formed a full factorial grid of 64 control conditions. Implementation details are provided in Appendix A.2.

3.5 Experimental Protocol

Training protocol Each control condition was trained with five random seeds. Models were optimised using Adam (Kingma & Ba, 2015) with a learning rate of 10^{-3} for at most 2,000 epochs. Training details are provided in Appendix A.3.

Evaluation metrics Functional organisation was characterised using unit-level task preferences measured under tracking and holding demands, summarised using two complementary measures: **(1) signed role alignment**, which measures whether hidden-unit task preferences are relatively more tracking-oriented in the left controller and more holding-oriented in the right controller, and **(2) functional separation**, defined by the Wasserstein distance between the two controller task-preference distributions. Behavioural performance was measured by **bimanual success rate (Bimanual SR)**, **tracking success rate (Tracking SR)**, and **holding success rate (Holding SR)**, and control efficiency by **mean control energy** and **energy per success**. Definitions are in Appendix B.

4 Results

4.1 Task assignment and training objectives interact to strengthen specialisation

Functional specialisation was strongest when controller-specific objectives were combined with fixed hand–task assignment. Across both organisation measures, either factor alone produced weaker organisation, indicating an interaction between objective structure and task assignment.

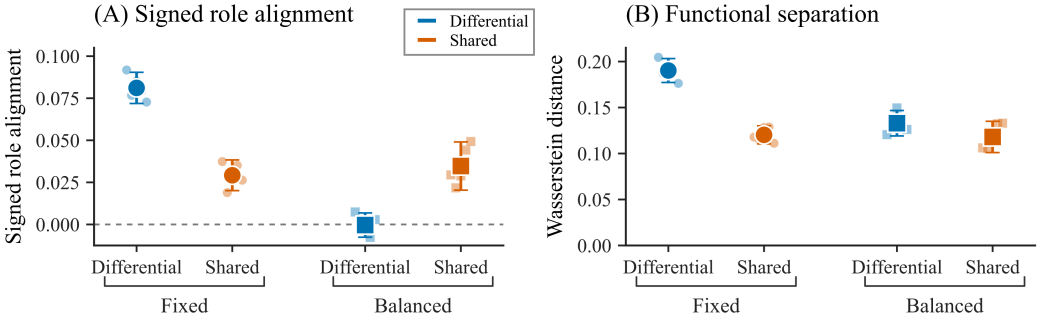

Figure 3: Functional specialisation across task assignments and training objectives. **(A)** Role alignment across training conditions, where larger positive values indicate stronger alignment with the assigned tracking and holding roles, near-zero values indicate little role-aligned preference, and negative values indicate reversed alignment. **(B)** Functional separation across training conditions, where larger values indicate greater separation between controller task-preference distributions. Large markers show five-seed means. Small markers show seed means across 64 control conditions, with horizontal jitter for visibility. Error bars denote 95% Student's t confidence intervals across seeds.

Signed role alignment was highest under Differential-Fixed training (0.081), compared with 0.029 for Shared-Fixed, -0.0004 for Differential-Balanced, and 0.035 for Shared-Balanced (Fig. 3A). Differential-Balanced therefore showed almost no task-aligned preference despite retaining controller-specific objectives. Functional separation showed the same overall pattern (Fig. 3B).

Differential-Fixed produced the largest separation (0.190), compared with 0.120 for Shared-Fixed, 0.133 for Differential-Balanced, and 0.118 for Shared-Balanced.

To test whether the effect of training objectives depended on task assignment, we computed an objective-by-assignment interaction contrast within each seed after averaging over the 64 control conditions defined in Section 3.5. The interaction was positive for both signed role alignment and functional separation (significant at $p < 0.05$). Thus, differential objectives strengthened functional specialisation more under fixed than balanced hand–task assignment, across both organisation measures. Full-grid results are provided in Appendix C.1.

4.2 Specialisation Accompanies Higher Performance and Control Efficiency

Differential-Fixed achieved the highest Bimanual SR (81.33%), exceeding all alternatives. The higher Bimanual SR of Differential-Fixed was reflected in both task components, with the highest tracking and holding success rates (92.31% and 87.08%) and the lowest errors (Table 1).

Table 1: Bimanual performance and control efficiency under the right-tracking/left-holding mapping. Values are mean ± standard deviation across five means, each averaged over 64 control conditions defined in Section 3.5. * denotes a significant difference from Differential-Fixed after Holm correction ($p_{\text{Holm}} < 0.05$). Arrows indicate the preferred direction of each metric.

Metric	Differential-Fixed	Shared-Fixed	Differential-Balanced	Shared-Balanced
Bimanual SR ↑	**81.33 ± 3.93%**	71.10 ± 6.69%*	28.03 ± 7.45%*	41.36 ± 4.92%*
Tracking SR ↑	**92.31 ± 1.88%**	89.28 ± 2.35%*	64.33 ± 4.17%*	84.38 ± 2.16%*
Holding SR ↑	**87.08 ± 4.67%**	76.77 ± 7.81%*	45.05 ± 10.03%*	46.82 ± 5.69%*
Trajectory error (mm) ↓	**13.45 ± 0.60**	14.00 ± 0.42*	20.29 ± 2.70*	15.12 ± 0.42*
Holding error (mm) ↓	**6.79 ± 0.60**	7.99 ± 0.63*	11.63 ± 0.80*	10.81 ± 0.34*
Mean control energy ↓	0.550 ± 0.135	0.683 ± 0.171*	**0.510 ± 0.093**	0.647 ± 0.120*
Energy per success ↓	**0.679 ± 0.172**	0.975 ± 0.291*	1.955 ± 0.712*	1.558 ± 0.168*

Differential-Fixed also maintained relatively low control energy (0.550), below both Shared conditions. Differential-Balanced used slightly less energy (0.510), although this difference was not significant and was accompanied by substantially lower success. When control expenditure was considered relative to successful behaviour, Differential-Fixed achieved the lowest energy per success (0.679), indicating the highest control efficiency among the four training conditions.

Across the 64 control conditions, Differential-Fixed also showed the highest median Bimanual SR and the lowest median energy per success, indicating that its performance and efficiency advantages were broadly reflected across the control-condition grid (Fig. 4A–B).

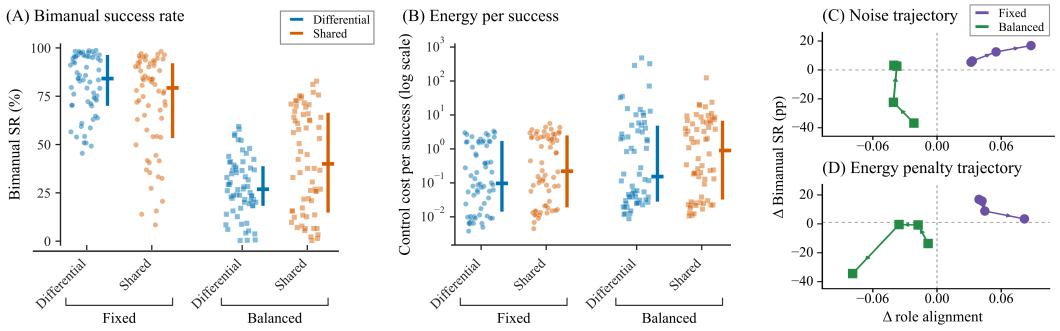

Figure 4: Bimanual performance and control efficiency across control conditions. **(A)** Bimanual SR under the canonical mapping. **(B)** Energy per success (log scale). Points show five-seed means for each of 64 control conditions. Horizontal and vertical bars indicate the median and interquartile range across conditions. **(C)** Differential–Shared differences in signed role alignment and Bimanual SR as **(C)** motor-noise gain or **(D)** energy penalty increases, averaging over the remaining factors. Positive Δ values indicate higher values under Differential training.

At the same time, behavioural gains were not proportional to specialisation strength (Fig. 4C–D). Along the motor-noise trajectory, larger Differential–Shared differences in signed role alignment were accompanied by larger Bimanual SR advantages under fixed assignment. Along the energy-penalty trajectory, the Bimanual SR advantage instead declined towards zero while the role-alignment difference remained positive. Taken together, the condition with the strongest task-aligned functional specialisation also achieved the highest performance and best control efficiency, but stronger specialisation did not consistently correspond to larger behavioural gains across control conditions. Full-grid results and statistical comparisons are provided in Appendices C.1 and C.2.

4.3 Differential Training Objectives Become More Advantageous Under Higher Motor Noise

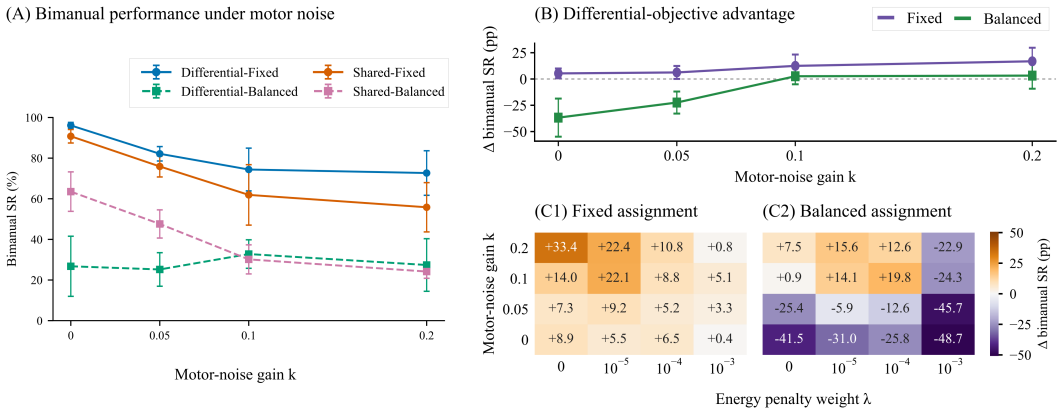

Figure 5: Bimanual performance under control constraints. **(A)** Success rate across motor-noise gain for four training conditions. **(B)** Differential–Shared Bimanual SR differences under fixed and balanced assignment. **(C)** Differential–Shared differences across noise and energy penalty under **(C1)** fixed and **(C2)** balanced assignment. Results average over undisplayed factors, with contrasts computed within matched conditions and seeds. A–B error bars show 95% Student's t confidence intervals across seeds. Positive differences favour Differential training.

Absolute performance shows that Differential-Fixed remains the highest-performing training condition across all tested motor-noise levels (Fig. 5A). As motor-noise gain increases, stronger signal-dependent perturbations make control more difficult, and Bimanual SR declines for most training conditions. Differential-Fixed nevertheless maintains comparatively high performance throughout the tested range. Shared-Fixed also remains substantially stronger than the balanced conditions, although its performance declines more sharply as motor-noise gain increases.

Given that Differential-Fixed achieved the best performance, we next examined how the performance of Differential and Shared training changes with motor-noise gain (Fig. 5B). Under fixed assignment, the Differential–Shared Bimanual SR advantage increased with motor-noise gain, reaching 17 percentage points at $k = 0.2$ when averaged across energy penalties. Under balanced assignment, Differential training performed substantially worse than Shared training at low motor-noise gain, but this disadvantage narrowed and approached zero at the two highest gain levels. Thus, at higher motor-noise gain, the relative performance benefit of differential training became more pronounced under both task-assignment schemes, with a consistent advantage under fixed assignment.

Energy penalty further modulated this effect (Fig. 5C). Differential training showed its largest advantages under weaker energy penalties, particularly when motor noise was high. Under fixed assignment, the largest difference reached 33.4 percentage points at $\lambda = 0$ and $k = 0.2$, whereas the advantage became small at stronger penalties. Under balanced assignment, Differential training exceeded Shared training mainly under moderate-to-high motor noise and weak-to-moderate energy penalties. Stronger energy penalties therefore reduced the additional performance benefit of differential training objectives. Taken together, the relative performance of differential objectives was more favourable at higher motor-noise gain, particularly under fixed task assignment. The comparatively high performance of Differential-Fixed at larger motor-noise gains indicates greater tolerance to motor-command-dependent execution noise.

4.4 COMMUNICATION DELAY INCREASES PERFORMANCE BUT NON-MONOTONICALLY ALTERS SPECIALISATION

Across four training conditions and control settings, Bimanual SR increased from 50.38% at $d = 0$ to 53.57%, 56.32%, and 61.55% at $d = 5$, 10, and 20 simulation steps (50, 100, and 200 ms), respectively. Functional specialisation showed a non-monotonic response to communication delay. Signed role alignment decreased from 0.0502 at $d = 0$ to 0.0312 and 0.0217 at $d = 5$ and $d = 10$, before returning to 0.0416 at $d = 20$. Functional separation showed a similar pattern, decreasing from 0.1533 to 0.1329 and 0.1280 before recovering to 0.1473 at $d = 20$. The largest reduction in both specialisation measures occurred at the delay of $d = 10$, whereas the longest delay produced values close to the no-delay condition. Because communication delay did not show a graded relationship with functional specialisation, it was treated as a categorical factor rather than a graded control axis in the trend analyses. The comparisons are reported in Appendix Table 6.

4.5 ROLE-SPECIFIC PERFORMANCE–FLEXIBILITY TRADE-OFF

To assess whether the performance advantage under trained hand–task roles transferred to altered task demands without retraining, we evaluated reversed hand–task mappings, symmetric bimanual tasks in which both hands performed tracking or holding, and perturbed-tracking tasks in which either hand tracked while resisting external perturbations (Fig. 6). Task definitions, evaluation procedures, and full results are provided in Appendix D.

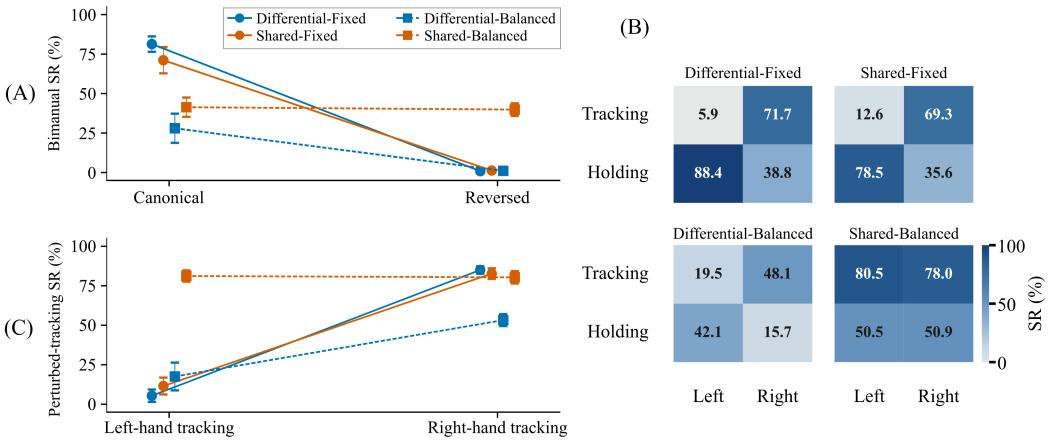

Figure 6: Hand-role flexibility across changing task demands without retraining. **(A)** Bimanual SR under canonical and reversed hand–task mappings. **(B)** Hand-specific tracking and holding success rates when both hands perform the same function. **(C)** Perturbed-tracking success rate when either hand tracks under external perturbation. Means are averaged over 64 control conditions. Panels A and C show 95% Student's t confidence intervals across five seeds.

Performance under altered hand–task roles differed substantially across training conditions (Fig. 6A–B). Differential-Fixed achieved 81.33% Bimanual SR under the canonical mapping but only 0.87% after role reversal, with tracking and holding competence concentrated in the trained hands. Shared-Fixed showed a similar dependence on trained hand roles. By comparison, Shared-Balanced retained similar Bimanual SR across canonical and reversed mappings (41.36% and 39.85%) and more even competence across the two hands. Differential-Balanced remained poor after role reversal, showing that balanced assignment alone was insufficient. Reversal reduced Bimanual SR by 80.46 percentage points for Differential-Fixed, compared with only 1.51 percentage points for Shared-Balanced. The same pattern appeared under perturbed tracking (Fig. 6C). Differential-Fixed achieved 85.01% success when tracking was assigned to the trained hand but only 5.39% with the opposite hand, whereas Shared-Balanced performed similarly with either hand (80.34% and 81.17%). The between-hand gap was therefore 79.62 percentage points for Differential-Fixed but only 0.83 percentage points for Shared-Balanced.

These results indicate that stronger role-specific specialisation was associated with higher performance when task demands matched the trained hand roles, but lower flexibility when those roles

changed. Shared-Balanced training showed the opposite pattern, with lower canonical-task performance but greater flexibility across hands and task demands.

5 Discussion and conclusion

Our main finding is that functional specialisation, a division of labour in which different controllers contribute according to their respective strengths, was associated with behavioural advantages in bilateral robotic control. Differential-Fixed, which developed the strongest functional specialisation, achieved the highest bimanual success, with the highest tracking and holding success rates and the lowest errors, together with the lowest energy per success under the trained tasks. Further, this advantage remained evident under challenging control conditions. Higher motor-noise gain increases the magnitude of execution noise for the same motor command, reducing command-to-output reliability, which can result from component wear, abnormal friction, and joint backlash in robotic systems (He et al., 2023; Kim, 2026). Beyond execution reliability, robotic systems face energy constraints, making both tolerance to unreliable motor execution and energy-efficient control important for maintaining performance and extending operational capacity. The model with the strongest functional specialisation, Differential-Fixed, remained the strongest performer as motor-noise gain increased, with its advantage becoming more favourable at higher noise levels, suggesting greater tolerance to unreliable motor execution. It also remained the strongest performer across energy penalties, although its advantage narrowed as energetic cost was weighted more strongly.

This behaviourally advantageous functional specialisation depended on how learning responsibility was organised across controllers and hand roles within a system in which each controller has a stronger routing gain to one hand. Controller-specific functional objectives produced the clearest task-aligned specialisation when paired with stable hand–task assignment, as in Differential-Fixed. Differential-Balanced training makes this interaction clear: controller-specific objectives alone were insufficient. Under fixed assignment, each controller was repeatedly associated with the same function through its more strongly routed hand (Arac et al., 2026), whereas balanced assignment varied these hand-specific roles and produced weaker task-aligned functional specialisation.

Biological motor systems balance functional specialisation with redundant control, allowing different components to develop complementary strengths while retaining some capacity to compensate when normal roles change (Serrien et al., 2006). Our results reveal a similar trade-off in bilateral robotic control. Differential-Fixed achieved the strongest role-specific performance and efficiency, but its competence was concentrated in the trained hand–task roles, leading to a large loss of performance when those roles were reversed or reassigned. Shared-Balanced showed the opposite pattern: lower performance under the trained mapping, but much greater interchangeability across hands and altered task demands. Functional specialisation therefore affected not only internal controller organisation but also how behavioural competence was distributed across the system. For robotic control, this suggests that stronger specialisation may be advantageous when task roles are stable and repeatedly reused, whereas greater redundancy may be preferable when hand roles must change or when the system needs to compensate for altered operating demands.

Limitations and future work Our findings are limited to the bilateral motor-control setting and task family studied here. Biological principles are used to motivate the separation of controller-level functional specialisation and stable hand-role preference, but our model does not attempt to reproduce the biological mechanisms underlying human handedness or hemispheric specialisation. Hand-role flexibility is evaluated only through reassignment within this setting, leaving broader generalisation across task families and embodiments outside the scope. Because contralateral routing was fixed across all training conditions, the results characterise learning responsibility within a bilateral architecture in which each controller has a stronger routing gain to one hand and a weaker gain to the other. Alternative controller-to-arm mappings are not examined. Because Fixed and Balanced conditions differed in the hand–task mappings encountered during training, the observed flexibility differences reflect training exposure together with objective structure and learned organisation. More generally, because training structure affects both functional organisation and behaviour, these experiments do not isolate whether specialisation itself mediates the observed behavioural advantages independently of the training structure that produced it.

Future work could test broader task families, simulation environments, and physical robots, and compare with human behavioural or neural data if extending the biological interpretation.

AI USE STATEMENT

Generative AI tools were used to refine and improve wording based on author-written drafts, with all AI-assisted text manually reviewed and further edited by the authors. AI was also used to support literature retrieval, including identifying potentially relevant work for related-work discussions and methodological references, with all cited papers read and verified by the authors. All references were managed by the authors by importing the original PDFs into Zotero, exporting the corresponding BibTeX entries, and manually checking each entry. No reference entry was directly generated by AI. AI tools assisted with code generation and editing, and all resulting code was reviewed by the authors. They were also used for tooling tasks, such as generating editable source SVG files for architecture diagrams, all of which were subsequently manually modified. No directly AI-generated figures were used in the paper. The authors take full responsibility for the final content of this work.

ETHICS STATEMENT

This work consists entirely of computational simulations and does not involve human participants, animal experiments, or the collection or use of personal data. The controllers are evaluated in a simulated environment, with no deployment on physical robots or testing involving people.

REPRODUCIBILITY STATEMENT

Experiments used Python 3.10.19, PyTorch 2.5.1 with CUDA 12.1, NumPy 2.2.5, and MotorNet 0.2.0. Models were trained on NVIDIA L40S-class GPUs. The complete sweep comprised 1,280 training runs and required approximately 704 GPU-hours. Code is provided in an anonymised repository at: anonymous.4open.science/r/code202609-671B.

REFERENCES

Ahmet Arac, Nicolas Y. H. Jeong Lee, and John W. Krakauer. Arm dominance is an emergent effect of practice executing complex trajectory shapes required by tools and objects. *Proceedings of the National Academy of Sciences*, 123(27):e2601569123, 2026. doi: 10.1073/pnas.2601569123.

Hao Ban and Kaiyi Ji. Fair resource allocation in multi-task learning. In *Proceedings of the 41st International Conference on Machine Learning*, pp. 2715–2731, 2024.

Marlene Behrmann, Nicholas M. Blauch, and David C. Plaut. The Emergence of Topography and Hemispheric Lateralization in High-Level Vision. *Quarterly Journal of Experimental Psychology*, 79(3):546–570, 2026. doi: 10.1177/17470218251395631.

Jacoba Brinkman and Henricus G. J. M. Kuypers. Cerebral control of contralateral and ipsilateral arm, hand and finger movements in the split-brain rhesus monkey. *Brain*, 96(4):653–674, 1973. doi: 10.1093/brain/96.4.653.

Ya-Ning Chang and Matthew A. Lambon Ralph. A unified neurocomputational bilateral model of spoken language production in healthy participants and recovery in poststroke aphasia. *Proceedings of the National Academy of Sciences*, 117(51):32779–32790, 2020. doi: 10.1073/pnas.2010193117.

Zitian Chen, Yikang Shen, Mingyu Ding, Zhenfang Chen, Hengshuang Zhao, Erik Learned-Miller, and Chuang Gan. Mod-squad: Designing mixtures of experts as modular multi-task learners. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11828–11837, 2023. doi: 10.1109/CVPR52729.2023.01138.

Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using RNN encoder–decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1724–1734, 2014. doi: 10.3115/v1/D14-1179.

Olivier Codol, Jonathan A. Michaels, Mehrdad Kashefi, J. Andrew Pruszynski, and Paul L. Gribble. MotorNet, a Python toolbox for controlling differentiable biomechanical effectors with artificial neural networks. *eLife*, 12:RP88591, 2024. doi: 10.7554/eLife.88591.

Chengyu Deng, Guanqi Chen, Yizhou Chen, Zejia Liu, Zhiwen Ruan, Guanhua Chen, and Jia Pan. Semantically structured mixture-of-experts for compositional robotic manipulation. In *Proceedings of Robotics: Science and Systems*, 2026. doi: 10.15607/RSS.2026.XXII.057.

Brooke Dexheimer, Robert Sainburg, Sydney Sharp, and Benjamin A. Philip. Roles of Handedness and Hemispheric Lateralization: Implications for Rehabilitation of the Central and Peripheral Nervous Systems: A Rapid Review. *The American Journal of Occupational Therapy*, 78(2): 7802180120, 2024. doi: 10.5014/ajot.2024.050398.

Chuntao Ding, Zhichao Lu, Shangguang Wang, Ran Cheng, and Vishnu N. Boddeti. Mitigating task interference in multi-task learning via explicit task routing with non-learnable primitives. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7756–7765, 2023. doi: 10.1109/CVPR52729.2023.00749.

Katharina Dobs, Julio Martinez, Alexander J. E. Kell, and Nancy Kanwisher. Brain-like functional specialization emerges spontaneously in deep neural networks. *Science Advances*, 8(11): eabl8913, 2022. doi: 10.1126/sciadv.abl8913.

William Fedus, Barret Zoph, and Noam Shazeer. Switch Transformers: Scaling to Trillion Parameter Models with Simple and Efficient Sparsity. *Journal of Machine Learning Research*, 23(120):1–39, 2022.

Antonia F. de C. Hamilton and Daniel M. Wolpert. Controlling the statistics of action: Obstacle avoidance. *Journal of Neurophysiology*, 87(5):2434–2440, 2002. doi: 10.1152/jn.2002.87.5.2434.

Christopher M. Harris and Daniel M. Wolpert. Signal-dependent noise determines motor planning. *Nature*, 394(6695):780–784, 1998. doi: 10.1038/29528.

Masahiko Haruno, Daniel M. Wolpert, and Mitsuo Kawato. Mosaic model for sensorimotor learning and control. *Neural Computation*, 13(10):2201–2220, 2001. doi: 10.1162/089976601750541778.

Yiming He, Jihong Chen, Xing Zhou, and Shifeng Huang. In-situ fault diagnosis for the harmonic reducer of industrial robots via multi-scale mixed convolutional neural networks. *Journal of Manufacturing Systems*, 66:233–247, 2023. doi: 10.1016/j.jmsy.2022.12.001.

Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979.

Wenbin Hu, Fernando Acero, Eleftherios Triantafyllidis, Zhaocheng Liu, and Zhibin Li. Modular neural network policies for learning in-flight object catching with a robot hand-arm system. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 944–951, 2023. doi: 10.1109/IROS55552.2023.10341463.

Peter J. Huber. Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35(1):73–101, 1964. doi: 10.1214/aoms/1177703732.

Pingcheng Jian, Easop Lee, Zachary Bell, Michael M. Zavlanos, and Boyuan Chen. Policy stitching: Learning transferable robot policies. In *Proceedings of the 7th Conference on Robot Learning*, pp. 3789–3808, 2023.

Heonkook Kim. Physics-guided residual learning and SHAP attribution for explainable robotic fault diagnosis. *Journal of Intelligent Manufacturing*, 2026. doi: 10.1007/s10845-026-02865-9.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations*, 2015.

Jingwen Li, Levin Kuhlmann, Jason Friedman, and Gideon Kowadlo. Communication in modular robotic motor control: Bilateral controllers under realistic constraints. *arXiv preprint arXiv:2608.13904*, 2026. doi: 10.48550/arXiv.2608.13904.

Zechu Li, Yufeng Jin, Daniel Ordonez-Apraez, Claudio Semini, Puze Liu, and Georgia Chalvatzaki. Morphologically symmetric reinforcement learning for ambidextrous bimanual manipulation. In *Proceedings of the 9th Conference on Robot Learning*, pp. 1953–1974, 2025.

Bo Liu, Yihao Feng, Peter Stone, and Qiang Liu. FAMO: Fast adaptive multitask optimization. In *Advances in Neural Information Processing Systems*, pp. 57226–57243, 2023. doi: 10.52202/075280-2500.

I-Chun Arthur Liu, Sicheng He, Daniel Seita, and Gaurav S. Sukhatme. VoxAct-B: Voxel-based acting and stabilizing policy for bimanual manipulation. In *Proceedings of the 8th Conference on Robot Learning*, pp. 4354–4370, 2025.

Gianpiero Liuzzi, Vanessa Hörniss, Máximo Zimerman, Christian Gerloff, and Friedhelm C. Hummel. Coordination of Uncoupled Bimanual Movements by Strictly Timed Interhemispheric Connectivity. *Journal of Neuroscience*, 31(25):9111–9117, 2011. doi: 10.1523/JNEUROSCI.0046-11.2011.

Alexandre Jehan Marcori, Pedro Henrique Martins Monteiro, and Victor Hugo Alves Okazaki. Changing handedness: What can we learn from preference shift studies? *Neuroscience & Biobehavioral Reviews*, 107:313–319, 2019. doi: 10.1016/j.neubiorev.2019.09.019.

Bernd-Ulrich Meyer, Simone Röricht, Helga Gräfin von Einsiedel, Frithjof Kruggel, and Adolph Weindl. Inhibitory and excitatory interhemispheric transfers between motor cortical areas in normal humans and patients with abnormalities of the corpus callosum. *Brain*, 118(2):429–440, 1995. doi: 10.1093/brain/118.2.429.

Jonathan A. Michaels, Stefan Schaffelhofer, Andres Agudelo-Toro, and Hansjörg Scherberger. A goal-driven modular neural network predicts parietofrontal neural dynamics during grasping. *Proceedings of the National Academy of Sciences*, 117(50):32124–32135, 2020. doi: 10.1073/pnas.2005087117.

Pratik K. Mutha, Kathleen Y. Haaland, and Robert L. Sainburg. The effects of brain lateralization on motor control and adaptation. *Journal of Motor Behavior*, 44(6):455–469, 2012. doi: 10.1080/00222895.2012.747482.

Aviv Navon, Aviv Shamsian, Idan Achituve, Haggai Maron, Kenji Kawaguchi, Gal Chechik, and Ethan Fetaya. Multi-task learning as a bargaining game. In *Proceedings of the 39th International Conference on Machine Learning*, pp. 16428–16446, 2022.

Lisa Pauwels and Jolien Gooijers. The role of the corpus callosum (micro)structure in bimanual coordination: A literature review update. *Journal of Motor Behavior*, 55(5):525–537, 2023. doi: 10.1080/00222895.2023.2221985.

Gabriel Peyré and Marco Cuturi. Computational optimal transport. *Foundations and Trends in Machine Learning*, 11(5–6):355–607, 2019. doi: 10.1561/2200000073.

Belinda Phipson and Gordon K. Smyth. Permutation p-values should never be zero: Calculating exact p-values when permutations are randomly drawn. *Statistical Applications in Genetics and Molecular Biology*, 9(1):Article 39, 2010. doi: 10.2202/1544-6115.1585.

Kenneth A. Provins. The specificity of motor skill and manual asymmetry: A review of the evidence and its implications. *Journal of Motor Behavior*, 29(2):183–192, 1997. doi: 10.1080/00222899709600832.

Jarrad Rinaldo, Jason Friedman, Levin Kuhlmann, and Gideon Kowadlo. Left/Right Brain, Human Motor Control and the Implications for Robotics. In Giuseppe Nicosia, Varun Ojha, Sven Giesselbach, M. Panos Pardalos, and Renato Umeton (eds.), *Machine Learning, Optimization, and Data Science*, volume 15510 of *Lecture Notes in Computer Science*, pp. 261–274. Springer Nature Switzerland, Cham, 2025. doi: 10.1007/978-3-031-82487-6_18.

Robert L. Sainburg. Handedness: Differential Specializations for Control of Trajectory and Position. *Exercise and Sport Sciences Reviews*, 33(4):206–213, 2005. doi: 10.1097/00003677-200510000-00010.

Luís F. Seoane. Optimality Pressures toward Lateralization of Complex Brain Functions. *Physical Review X*, 13(3):031028, 2023. doi: 10.1103/PhysRevX.13.031028.

Deborah J. Serrien, Richard B. Ivry, and Stephan P. Swinnen. Dynamics of hemispheric specialization and integration in the context of motor control. *Nature Reviews Neuroscience*, 7(2):160–166, 2006. doi: 10.1038/nrn1849.

Noam Shazeer, Azalia Mirhoseini, Krzysztof Maziarz, Andy Davis, Quoc Le, Geoffrey Hinton, and Jeff Dean. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. In *International Conference on Learning Representations*, 2017.

Stephan P. Swinnen. Intermanual coordination: From behavioural principles to neural-network interactions. *Nature Reviews Neuroscience*, 3(5):348–359, 2002. doi: 10.1038/nrn807.

Emanuel Todorov and Michael I. Jordan. Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, 5(11):1226–1235, 2002. doi: 10.1038/nn963.

Daniel M. Wolpert and Mitsuo Kawato. Multiple paired forward and inverse models for motor control. *Neural Networks*, 11(7-8):1317–1329, 1998. doi: 10.1016/S0893-6080(98)00066-5.

Elizabeth J. Woytowicz, Kelly P. Westlake, Jill Whitall, and Robert L. Sainburg. Handedness results from complementary hemispheric dominance, not global hemispheric dominance: Evidence from mechanically coupled bilateral movements. *Journal of Neurophysiology*, 120(2):729–740, 2018. doi: 10.1152/jn.00878.2017.

Guangyu Robert Yang, Madhura R. Joglekar, H. Francis Song, William T. Newsome, and Xiao-Jing Wang. Task representations in neural networks trained to perform many cognitive tasks. *Nature Neuroscience*, 22(2):297–306, 2019. doi: 10.1038/s41593-018-0310-2.

Tengbo Yu, Guanxing Lu, Zaijia Yang, Haoyuan Deng, Season Si Chen, Jiwen Lu, Wenbo Ding, Guoqiang Hu, Yansong Tang, and Ziwei Wang. ManiGaussian++: General robotic bimanual manipulation with hierarchical Gaussian world model. In *2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 12232–12239, 2025. doi: 10.1109/IROS60139.2025.11246564.

Yu Zhang and Qiang Yang. A survey on multi-task learning. *IEEE Transactions on Knowledge and Data Engineering*, 34(12):5586–5609, 2022. doi: 10.1109/TKDE.2021.3070203.

A EXPERIMENTAL SETUP AND TRAINING

A.1 MUSCULOSKELETAL ENVIRONMENT AND TASK DEFINITIONS

Simulations used MotorNet 0.2.0 with the default parameters of RigidTendonArm26 and RigidTendonHillMuscle, comprising shoulder flexor/extensor, elbow flexor/extensor, and biarticular flexor/extensor muscles (Codol et al., 2024). The simulation step was $\Delta t = 0.01$ s and each episode lasted $T = 100$ steps. Shoulder origins were separated by 0.4 m, with the left arm mirrored horizontally in the workspace. Arms started from shoulder and elbow angles of $45°$ and $90°$, respectively.

For the tracking task, let $\mathbf{r}_0$ denote the initial endpoint of the tracking hand. The target trajectory was

$$\mathbf{r}_t = \mathbf{r}_0 + sa\sin(2\pi ft\Delta t)\mathbf{e}_y, \tag{4}$$

where $a \sim \mathcal{U}(0.04, 0.06)$ m, $f \sim \mathcal{U}(0.75, 1.25)$ Hz, and $s \in \{-1, +1\}$ was sampled with equal probability. The holding target remained at the initial endpoint of the holding hand.

The holding hand received a constant Cartesian force pulse with onset sampled uniformly from steps 5–10, duration from 10–15 steps, magnitude from 0.5–1.5 N, and direction uniformly from $[0, 2\pi)$. The tracking hand received no external force pulse in the main task.

A.2 CONTROLLER ARCHITECTURE AND CONTROL CONSTRAINTS

Let x_t denote the shared 46-dimensional sensory observation. It comprised proprioceptive muscle states, hand endpoint positions, target-relative features, and bilateral spatial features describing the relationship between the two hands. The recurrent updates were

$$h_t^L = \mathrm{GRU}_L([x_t, C_{R\to L}\bar{h}_t^R], h_{t-1}^L), \tag{5}$$
$$h_t^R = \mathrm{GRU}_R([x_t, C_{L\to R}\bar{h}_t^L], h_{t-1}^R). \tag{6}$$

Here, $C_{R\to L}$ and $C_{L\to R}$ are learned communication projections. For a state from controller $j \in \{L, R\}$, the available communicated state was $\bar{h}_t^j = h_{t-d}^j$ for $d > 0$; for $d = 0$, the latest causally available state $\bar{h}_t^j = h_{t-1}^j$ was used. The tested positive delays of 5, 10, and 20 simulation steps therefore corresponded to 50, 100, and 200 ms.

To implement the fixed 80:20 contralateral-to-ipsilateral routing, let $p = 0.8$, $c = \sqrt{2/(p^2 + (1-p)^2)}$, and let $\mathbf{1}_6$ denote a six-dimensional vector of ones. With the first six motor components corresponding to the left arm and the remaining six to the right arm,

$$g_L = c[(1-p)\mathbf{1}_6, p\mathbf{1}_6], \qquad g_R = c[p\mathbf{1}_6, (1-p)\mathbf{1}_6]. \tag{7}$$

The deterministic motor command and its signal-dependent noisy execution were

$$u_t = \sigma(g_L \odot W_L h_t^L + g_R \odot W_R h_t^R + b), \tag{8}$$
$$\widetilde{u}_t = \mathrm{clip}_{[0,1]}\left(u_t \odot (1 + k\epsilon_t)\right), \qquad \epsilon_t \sim \mathcal{N}(0, I_{12}), \tag{9}$$

with fixed motor bias $b = 0.1$. The routing gains scaled the learned readout outputs and therefore did not impose a fixed fraction of executed control from either controller.

Communication delay d and motor-noise gain k acted directly on the forward control process defined above, whereas the energy-penalty weight λ entered the training objectives in Appendix A.3.

A.3 TRAINING PROCEDURE AND OBJECTIVES

Training configuration. Models were trained with Adam (Kingma & Ba, 2015) using a learning rate of 10^{-3} and a batch size of 128. Each epoch comprised one newly sampled batch of one-second

episodes and one parameter update. Training ran for at most 2,000 epochs. Five random seeds (0–4) were used for each control condition. Training used validation-based early stopping, and the checkpoint with the best validation performance over the hand–task mappings used during training was retained for evaluation.

Loss definitions and gradient assignment. For trial $b \in \{1, \dots, B\}$ and time step $t \in \{1, \dots, T\}$, let e_{bt}^T and e_{bt}^H denote the Euclidean endpoint errors of the tracking and holding hands, respectively. Here, B is the batch size and T is the number of simulation steps per episode. Both errors were penalised using the Huber loss $\rho_\delta(\cdot)$ with transition $\delta = 0.01$ m (Huber, 1964). Let m_{bt} be an indicator equal to one from perturbation onset to the end of the episode and zero otherwise, and let $u_{bt} \in \mathbb{R}^{12}$ denote the deterministic motor command before signal-dependent noise and clipping. The functional losses and motor-output penalty were

$$L_T = \frac{1}{BT} \sum_{b,t} \rho_\delta(e_{bt}^T), \tag{10}$$

$$L_H = \frac{\sum_{b,t} m_{bt} \rho_\delta(e_{bt}^H)}{\sum_{b,t} m_{bt}}, \tag{11}$$

$$E = \frac{1}{BT} \sum_{b,t} \|u_{bt}\|_2^2. \tag{12}$$

Thus, tracking loss was evaluated over the full episode, whereas holding loss was evaluated from perturbation onset through recovery. The energy term used the deterministic motor commands before signal-dependent noise and clipping.

The controller-owned parameter groups were

$$\theta_L = \{\mathrm{GRU}_L, W_L, C_{R\to L}\}, \qquad \theta_R = \{\mathrm{GRU}_R, W_R, C_{L\to R}\}. \tag{13}$$

Under Differential training, the gradients were

$$g_{\theta_L} = \nabla_{\theta_L}(L_T + \lambda E), \qquad g_{\theta_R} = \nabla_{\theta_R}(L_H + \lambda E). \tag{14}$$

Table 2: Training implementation for the learning-responsibility conditions. Canonical denotes right-hand tracking and left-hand holding. Reversed denotes left-hand tracking and right-hand holding.

Condition	Batch composition	Training objective
Differential-Fixed	128 canonical	$\theta_L : L_T + \lambda E, \quad \theta_R : L_H + \lambda E$
Differential-Balanced	64 canonical + 64 reversed	$\theta_L : L_T + \lambda E, \quad \theta_R : L_H + \lambda E$
Shared-Fixed	128 canonical	$\theta_L, \theta_R : L_T + L_H + \lambda E$
Shared-Balanced	64 canonical + 64 reversed	$\theta_L, \theta_R : L_T + L_H + \lambda E$

Under Differential-Balanced training, controller-level functional responsibilities remained fixed when the hand–task mapping was reversed. Gradient assignment affected parameter updates only. Both controllers, arms, and communication pathways remained active in the forward computation.

B Evaluation metrics

B.1 Functional specialisation metrics

For the controlled functional-specialisation assessment, each trained model was evaluated under the canonical hand–task mapping without parameter updates. Execution noise was set to zero to isolate responses to controlled task demands. Tracking frequency took values $\{0.75, 1.0, 1.25\}$ Hz and holding force magnitude took values $\{0.5, 1.0, 1.5\}$ N, giving a 3×3 factorial design. Tracking amplitude was fixed at 0.05 m, and each demand combination used 64 trials.

Unit-level task preference. For hidden unit i in trial b, activity was summarised by its root mean square over the episode,

$$a_{ib} = \sqrt{\frac{1}{T}\sum_{t=1}^{T} h_{ibt}^2}. \tag{15}$$

Within each evaluation stream, unit activity, tracking frequency, and holding force were standardised. Ordinary least squares fitted

$$z(a_{ib}) = \beta_{0i} + \beta_{Ti} z(f_b) + \beta_{Hi} z(F_b) + \beta_{THi} z(f_b) z(F_b) + \varepsilon_{ib}. \tag{16}$$

Adapting the fractional task-selectivity formulation of Yang et al. (2019), relative preference between the two functional demands was defined as

$$p_i = \frac{|\beta_{Ti}| - |\beta_{Hi}|}{\max(|\beta_{Ti}| + |\beta_{Hi}|, 10^{-12})}. \tag{17}$$

Positive values indicate greater sensitivity to tracking frequency and negative values greater sensitivity to holding force. All 64 units per controller were included.

Signed role alignment. Let $\bar{p}_L$ and $\bar{p}_R$ denote the mean unit preferences of the left and right controllers. Signed role alignment was

$$A_{\text{role}} = \frac{\bar{p}_L - \bar{p}_R}{2}. \tag{18}$$

Controller identities were kept fixed when computing role alignment and were not swapped based on their learned preferences. Positive values indicate relatively greater tracking preference in the left controller.

Functional separation. Functional separation was defined as the one-dimensional Wasserstein distance between the two empirical preference distributions (Peyré & Cuturi, 2019):

$$D_W = \frac{1}{64}\sum_{i=1}^{64} |p_{(i)}^L - p_{(i)}^R|, \tag{19}$$

where parentheses indicate separately sorted unit preferences. Unlike signed role alignment, this measure is invariant to exchanging controller labels and quantifies the magnitude of between-controller differentiation irrespective of direction.

These two functional-specialisation metrics quantify relative sensitivity to the manipulated tracking and holding demands under the canonical hand–task mapping.

B.2 Behavioural performance metrics

For the performance assessment, canonical and reversed mappings were each evaluated over 128 trials sampled from the same task distributions as training. Model parameters were frozen, and execution-noise gain was set to the model's training value k. For each control condition and training seed, model families were evaluated on the same set of sampled trials.

Tracking and holding errors. For trial b, the whole-episode mean tracking error and the pulse-and-recovery mean holding error were

$$\bar{e}_b^T = \frac{1}{T}\sum_t e_{bt}^T, \tag{20}$$

$$\bar{e}_b^H = \frac{\sum_t m_{bt} e_{bt}^H}{\sum_t m_{bt}}, \tag{21}$$

where m_{bt} selects time steps from perturbation onset through recovery. Reported trajectory and holding errors were the corresponding means across N evaluation trials:

$$\mathrm{Err}_T = \frac{1}{N}\sum_{b=1}^{N} \bar{e}_b^T, \qquad \mathrm{Err}_H = \frac{1}{N}\sum_{b=1}^{N} \bar{e}_b^H. \tag{22}$$

Success rates. Let $\mathbb{I}[\cdot]$ denote the indicator function. Tracking SR, Holding SR, and Bimanual SR were

$$S_T = \frac{1}{N}\sum_{b=1}^{N} \mathbb{I}[\bar{e}_b^T \le 0.02], \tag{23}$$

$$S_H = \frac{1}{N}\sum_{b=1}^{N} \mathbb{I}[\bar{e}_b^H \le 0.01], \tag{24}$$

$$S_B = \frac{1}{N}\sum_{b=1}^{N} \mathbb{I}[\bar{e}_b^T \le 0.02 \wedge \bar{e}_b^H \le 0.01]. \tag{25}$$

Bimanual SR therefore required both task criteria to be satisfied in the same trial. For symmetric evaluations, it analogously required both hands to satisfy the relevant tracking or holding criterion in the same trial. All success criteria were based on mean error.

B.3 Control efficiency metrics

Mean control energy. Mean control energy was computed from the deterministic motor commands u_{bt} defined in Appendix A.3, before signal-dependent noise and clipping, and included both successful and failed trials:

$$\overline{E} = \frac{1}{NT}\sum_{b=1}^{N}\sum_{t=1}^{T} \|u_{bt}\|_2^2. \tag{26}$$

Energy per success. Energy per success was defined as $\overline{E}/S_B$, where the numerator and denominator were averaged over the same analysis group before forming the ratio. It therefore represents control expenditure relative to success rate.

C Main task results and statistical analysis

C.1 Full-grid performance and functional organisation

Table 3 summarises all 64 constraint cells for each training condition. The cells comprise four communication delays, four energy-penalty weights, and four motor-noise gains. Each entry was calculated by first averaging the 64 cells within each training seed and then summarising the five resulting seed means. Bimanual SR uses each model's native execution-noise gain. Functional separation and signed role alignment use the noise-free controlled-demand evaluation defined in Appendix B.1.

Table 3: Performance and controller organisation across the complete constraint grid. Values are mean [95% confidence interval] across five seed-level grid means. Each seed-level value averages the same 64 constraint settings. Bimanual SR is reported as a percentage.

Training condition	Bimanual SR	Signed role alignment	Functional separation
Differential-Fixed	$81.33\ [76.45, 86.20]$	$0.0811\ [0.0719, 0.0904]$	$0.190\ [0.177, 0.203]$
Shared-Fixed	$71.10\ [62.80, 79.40]$	$0.0292\ [0.0201, 0.0384]$	$0.120\ [0.110, 0.130]$
Differential-Balanced	$28.03\ [18.78, 37.28]$	$-0.0004\ [-0.0076, 0.0068]$	$0.133\ [0.119, 0.147]$
Shared-Balanced	$41.36\ [35.25, 47.47]$	$0.0347\ [0.0204, 0.0490]$	$0.118\ [0.101, 0.135]$

Objective-by-assignment interactions. For each organisation metric, the interaction contrast was computed within each seed after averaging over the 64 constraint settings: $(\mathrm{Differential}-\mathrm{Fixed}-\mathrm{Shared}-\mathrm{Fixed})-(\mathrm{Differential}-\mathrm{Balanced}-\mathrm{Shared}-\mathrm{Balanced})$. The five seed-level contrasts were tested against zero using two-sided one-sample t tests ($df = 4$). Holm correction was applied jointly across the two organisation metrics. Both interactions remained significant after correction (signed role alignment: $p_{\mathrm{H}} = 3.96 \times 10^{-5}$; functional separation: $p_{\mathrm{H}} = 0.00295$).

Functional separation is non-negative and label-invariant, whereas signed role alignment preserves the anatomical controller ordering. A larger distribution distance can therefore coexist with little directional role alignment.

Supplementary controller lesion analysis We performed lesions as evaluation-time interventions on trained models without retraining or modifying learned weights. For the lesioned controller, both its outgoing communication signal and motor contribution were multiplied by a retained gain γ. For the left controller, $\tilde{\mathbf{z}}_{L\to R,t} = \gamma_L C_{L\to R}\bar{\mathbf{h}}_t^L$ and $\tilde{\mathbf{c}}_{L,t} = \gamma_L\left(\mathbf{g}_L \odot W_L\mathbf{h}_t^L\right)$, with the corresponding operation applied to the right controller. A gain of $\gamma = 1$ denotes the intact model and $\gamma = 0$ removes all outgoing influence from the lesioned controller while leaving its internal recurrent state and incoming signals unchanged. We therefore refer to $\gamma = 0$ as a whole-module functional lesion.

Lesions were evaluated under the canonical right-tracking/left-holding mapping using $\gamma \in \{0, 0.95, 0.975, 0.99\}$. For each checkpoint, intact, left-lesion, and right-lesion evaluations used matched trials and motor-noise realisations. Within each seed, outcomes were averaged equally over the 64 delay–energy–noise conditions and then summarised across five seeds. We report changes relative to the intact model as $\Delta = \mathrm{lesion} - \mathrm{intact}$.

Table 4: Graded controller lesion effects on canonical-task success rates. Values are mean ± SD across five seed-level means, with each seed averaged equally over the 64 control settings. Changes are calculated as lesion minus intact in percentage points. Retained gain γ scales the lesioned controller's outgoing influence, and negative values indicate reduced success after lesion.

		Δ Tracking SR (pp)				Δ Holding SR (pp)			
Model	Lesioned controller	$\gamma = 0$	$\gamma = 0.95$	$\gamma = 0.975$	$\gamma = 0.99$	$\gamma = 0$	$\gamma = 0.95$	$\gamma = 0.975$	$\gamma = 0.99$
Differential-Fixed	Left	-92.31 ± 1.88	-19.87 ± 8.33	-6.17 ± 3.01	-0.89 ± 0.91	-86.13 ± 4.26	-7.38 ± 1.36	-1.43 ± 0.71	-0.30 ± 0.19
	Right	-91.57 ± 1.70	-10.78 ± 6.14	-3.41 ± 2.77	-0.24 ± 0.23	-87.08 ± 4.67	-19.94 ± 2.62	-5.74 ± 1.67	-0.97 ± 0.63
Shared-Fixed	Left	-89.28 ± 2.35	-6.14 ± 2.84	-0.92 ± 0.71	-0.12 ± 0.32	-66.03 ± 4.67	-0.14 ± 0.24	$+0.02 \pm 0.23$	$+0.08 \pm 0.19$
	Right	-86.33 ± 2.83	-0.48 ± 0.45	-0.05 ± 0.25	$+0.08 \pm 0.16$	-76.77 ± 7.81	-5.59 ± 1.36	-1.63 ± 0.43	-0.45 ± 0.25
Differential-Balanced	Left	-64.33 ± 4.17	-40.68 ± 7.19	-26.06 ± 3.91	-10.50 ± 1.97	-45.05 ± 10.03	-36.38 ± 9.11	-27.51 ± 8.24	-12.93 ± 3.16
	Right	-64.33 ± 4.17	-37.72 ± 8.99	-23.25 ± 4.44	-7.68 ± 0.92	-45.05 ± 10.03	-41.62 ± 9.36	-34.99 ± 9.75	-19.78 ± 4.87
Shared-Balanced	Left	-84.38 ± 2.16	-3.25 ± 1.13	-0.43 ± 0.40	$+0.10 \pm 0.22$	-45.53 ± 6.05	-0.68 ± 0.93	-0.18 ± 0.53	-0.08 ± 0.24
	Right	-76.86 ± 5.54	-0.68 ± 0.45	-0.21 ± 0.24	-0.07 ± 0.07	-46.82 ± 5.69	-9.09 ± 3.38	-3.60 ± 1.67	-0.92 ± 0.47

Complete lesions ($\gamma = 0$) drove both task components to near-floor performance, showing that successful bimanual control depended on contributions from both controllers. Under partial attenuation, the direction of impairment was more selective: left-controller attenuation tended to reduce tracking more strongly, whereas right-controller attenuation reduced holding more strongly. Because controller identity is coupled to the primarily controlled hand by the fixed routing architecture, these directional effects reflect both controller–hand organisation and functional dependence. Lesion magnitude was therefore not treated as an independent measure of functional specialisation.

C.2 TASK PERFORMANCE COMPARISONS

The primary inferential question was whether differences between training conditions were consistently expressed across the 64 prespecified constraint settings. Accordingly, the five training seeds were first averaged within each constraint cell to obtain one estimate per training condition and setting, and the resulting 64 settings were treated as matched experimental conditions. Differential-Fixed was then compared with each alternative across the resulting 64 matched settings using two-sided paired Monte Carlo permutation tests.

Under the null hypothesis, condition labels were exchanged within each matched constraint setting. The test statistic was the difference between the two conditions after equal weighting of the 64 settings. Thus, the analysis tests for a systematic performance difference across the fixed constraint grid. For energy per success, the ratio was recomputed from the corresponding permuted energy and success-rate quantities after each label exchange.

Each comparison used 10^6 permutations with random seed 20260918. Monte Carlo probabilities used the plus-one correction $p = (1 + N_{\text{extreme}})/(10^6 + 1)$ (Phipson & Smyth, 2010). Holm correction was applied jointly to all 21 comparisons across seven metrics and three alternatives (Holm, 1979). Table 5 reports the Holm-adjusted results for all 21 canonical performance comparisons.

Table 5: Holm-adjusted p-values from paired Monte Carlo permutation tests for Differential-Fixed versus each alternative across the 64 prespecified constraint settings. Tests assess systematic condition differences across the fixed constraint grid and are two-sided. Correction was applied jointly across all 21 comparisons.

Metric	Shared-Fixed	Differential-Balanced	Shared-Balanced
Bimanual SR	< 0.001	< 0.001	< 0.001
Tracking SR	0.001	< 0.001	< 0.001
Holding SR	< 0.001	< 0.001	< 0.001
Trajectory error	0.004	< 0.001	< 0.001
Holding error	< 0.001	< 0.001	< 0.001
Mean control energy	< 0.001	0.162	< 0.001
Energy per success	< 0.001	< 0.001	< 0.001

The mean-control-energy difference between Differential-Fixed and Differential-Balanced was not significant after correction ($p_{\text{Holm}} = 0.162$). All other comparisons remained significant after Holm correction.

C.3 COMMUNICATION-DELAY RESULTS

To estimate the effect of communication delay, outcomes were averaged within each seed over training condition, energy penalty, and motor-noise gain separately for each delay level. Reported means were then calculated across five seeds. Each nonzero delay was compared with $d = 0$ using two-sided paired t tests for Bimanual SR, signed role alignment, and functional separation. Confidence intervals were calculated across five paired seed-level differences, and Holm correction was applied across the nine comparisons. Table 6 reports the resulting comparisons.

To examine whether the relative performance of Differential and Shared training varied across control regimes, Table 7 separates the Differential $-$ Shared Bimanual SR contrasts by communication delay, motor-noise gain, and task assignment. Within each combination, the four energy-penalty settings were averaged within each seed. The stratified contrasts varied substantially across control regimes and did not show a consistent monotonic relationship with communication delay.

D HAND-ROLE FLEXIBILITY ACROSS TASK CONFIGURATIONS

D.1 TASK DEFINITIONS AND EVALUATION PROTOCOL

Six task configurations were used to assess hand-role flexibility without parameter updates. M1 used the canonical mapping from the main task, whereas M2 reversed the hand roles. S1 and S2 tested

Table 6: Performance and functional specialisation by communication-delay level. Differences compare each nonzero delay with $d = 0$. Bimanual SR differences are reported in percentage points. p_{H} denotes Holm-adjusted p-values.

Metric	Delay d	Mean	Difference from $d = 0$ [95% CI]	p_{H}
Bimanual SR (%)	0	50.38	Reference	–
	5	53.57	$+3.19$ [-2.86, 9.24]	0.434
	10	56.32	$+5.94$ [4.55, 7.32]	0.003
	20	61.55	$+11.16$ [5.53, 16.80]	0.037
Signed role alignment	0	0.0502	Reference	–
	5	0.0312	-0.0190 [-0.0337, -0.0043]	0.116
	10	0.0217	-0.0285 [-0.0397, -0.0173]	0.017
	20	0.0416	-0.0086 [-0.0217, 0.0045]	0.427
Functional separation	0	0.1533	Reference	–
	5	0.1329	-0.0204 [-0.0390, -0.0017]	0.156
	10	0.1280	-0.0253 [-0.0393, -0.0114]	0.044
	20	0.1473	-0.0060 [-0.0189, 0.0070]	0.434

Table 7: Delay-stratified Differential − Shared Bimanual SR contrasts in percentage points.

Delay (steps)	Noise k	Fixed assignment ΔSR [95% CI]	Balanced assignment ΔSR [95% CI]
0	0	6.8 [1.0, 12.5]	-30.4 [-54.6, -6.1]
0	0.05	13.0 [-11.0, 37.0]	-22.1 [-45.1, 0.9]
0	0.1	16.5 [9.3, 23.8]	-5.1 [-17.5, 7.2]
0	0.2	20.4 [4.7, 36.2]	4.3 [-18.3, 26.9]
5	0	1.6 [-0.8, 4.1]	-29.7 [-52.2, -7.2]
5	0.05	2.1 [-7.8, 11.9]	-19.2 [-34.7, -3.7]
5	0.1	15.4 [4.0, 26.8]	7.5 [-6.4, 21.3]
5	0.2	30.7 [4.7, 56.8]	11.1 [-5.6, 27.7]
10	0	5.2 [-5.0, 15.4]	-38.9 [-52.8, -25.1]
10	0.05	7.2 [-2.9, 17.4]	-22.5 [-35.5, -9.6]
10	0.1	4.6 [-7.0, 16.3]	3.3 [-10.6, 17.2]
10	0.2	12.0 [-5.4, 29.3]	0.9 [-8.9, 10.6]
20	0	7.7 [-5.4, 20.9]	-48.0 [-74.5, -21.4]
20	0.05	2.7 [-9.9, 15.2]	-25.8 [-32.8, -18.8]
20	0.1	13.4 [-4.5, 31.3]	4.9 [-9.9, 19.7]
20	0.2	4.3 [-7.4, 16.0]	-3.3 [-17.2, 10.6]

symmetric tracking and holding demands, while P1 and P2 tested perturbed tracking with either hand active, as defined in Table 8.

Unless otherwise specified, target and perturbation distributions followed Appendix A.1. In S1, target amplitude, frequency, and movement direction were sampled independently for the two hands. In S2, pulse directions were sampled independently, while pulse magnitude and timing were shared within each trial. In P1 and P2, the inactive hand retained a stationary target without a pulse and was excluded from the success criterion. Each evaluation used 128 trials. For each evaluation and control condition, the four training conditions were assessed on the same sampled trials. Model parameters remained frozen, and execution-noise gain was set to the model's training value k.

Table 8: Definitions of the six hand–task configurations. Success criteria follow the behavioural metrics defined in Appendix B.2.

ID	Task demands	Success criterion
M1	Right hand tracks, left hand holds under a pulse (main task)	Tracking whole-episode mean error $\leq$ 20 mm and holding pulse-and-recovery mean error $\leq$ 10 mm in the same trial
M2	Left hand tracks, right hand holds under a pulse	Tracking whole-episode mean error $\leq$ 20 mm and holding pulse-and-recovery mean error $\leq$ 10 mm in the same trial
S1	Both hands track independent moving targets	Both hands' whole-episode mean tracking errors $\leq$ 20 mm
S2	Both hands hold fixed targets under external pulses	Both hands' pulse-and-recovery mean holding errors $\leq$ 10 mm
P1	Left hand tracks and receives a pulse	Left-hand mean tracking error $\leq$ 20 mm over both the full episode and pulse-and-recovery interval
P2	Right hand tracks and receives a pulse	Right-hand mean tracking error $\leq$ 20 mm over both the full episode and pulse-and-recovery interval

D.2 Performance across hand–task configurations

For each evaluation, the 64 constraint settings were first averaged within each training seed. Table 9 reports the mean and 95% Student's t confidence interval across the five resulting seed-level grid means.

Table 9: Performance across the six hand–task configurations. Values are success rates in percent, reported as mean [95% confidence interval] across five seed-level grid means.

ID	Metric	Differential-Fixed	Shared-Fixed	Differential-Balanced	Shared-Balanced
M1	Bimanual SR	81.33 [76.45, 86.20]	71.10 [62.80, 79.40]	28.03 [18.78, 37.28]	41.36 [35.25, 47.47]
M2	Bimanual SR	0.87 [−0.41, 2.14]	1.28 [0.34, 2.22]	1.07 [0.70, 1.45]	39.85 [35.87, 43.82]
S1	Bimanual SR	4.89 [1.39, 8.39]	9.78 [5.21, 14.34]	11.65 [5.90, 17.41]	63.52 [58.81, 68.22]
S1	Left-hand SR	5.94 [1.72, 10.17]	12.65 [7.07, 18.23]	19.51 [11.28, 27.74]	80.48 [76.27, 84.68]
S1	Right-hand SR	71.68 [58.84, 84.53]	69.34 [62.89, 75.80]	48.11 [44.02, 52.21]	78.01 [74.43, 81.59]
S2	Bimanual SR	36.67 [25.80, 47.54]	32.59 [29.04, 36.13]	9.70 [6.68, 12.71]	36.05 [27.99, 44.11]
S2	Left-hand SR	88.44 [83.05, 93.82]	78.50 [68.88, 88.12]	42.10 [30.91, 53.30]	50.52 [41.96, 59.08]
S2	Right-hand SR	38.81 [26.42, 51.19]	35.55 [29.80, 41.31]	15.70 [12.06, 19.33]	50.86 [45.02, 56.70]
P1	Perturbed-tracking SR	5.39 [1.46, 9.31]	11.53 [6.19, 16.87]	17.57 [8.77, 26.36]	81.17 [77.61, 84.74]
P2	Perturbed-tracking SR	85.01 [82.65, 87.38]	82.71 [79.36, 86.05]	53.26 [49.47, 57.04]	80.34 [76.45, 84.23]

The hand-specific rates clarify the source of performance differences in the symmetric tasks. The hand-specific rates distinguish bilateral task success from competence of the individual hands. In particular, low Bimanual SR can coexist with high success in one hand when performance is strongly asymmetric across hands.

D.3 Statistical tests of hand-role asymmetry

Four contrasts were used to quantify hand-role asymmetry within each training condition: (1) M1 − M2, (2) right − left tracking success in S1, (3) left − right holding success in S2, and (4) P2 − P1. Positive values indicate better performance in the canonical right-tracking/left-holding direction.

For each contrast, differences were first calculated within matched constraint settings and training seeds, then averaged over the 64 settings within each seed. The five resulting seed-level contrasts were tested against zero using two-sided one-sample t tests. Reported confidence intervals are 95% Student's t intervals across the five seed-level contrasts. Holm correction was applied jointly across the 16 tests.

Table 10: Within-condition hand-role asymmetry contrasts. Differences are reported in percentage points as mean [95% confidence interval]. p_{H} denotes the Holm-adjusted p-value across the 16 prespecified tests.

Contrast	Training condition	Difference [95% CI]	p_{H}
M1 − M2	Differential-Fixed	80.46 [75.56, 85.36]	< 0.001
	Shared-Fixed	69.82 [60.83, 78.80]	< 0.001
	Differential-Balanced	26.96 [17.67, 36.25]	0.009
	Shared-Balanced	1.51 [−3.29, 6.32]	1.000
S1 right − left	Differential-Fixed	65.74 [52.42, 79.06]	0.002
	Shared-Fixed	56.69 [47.25, 66.13]	< 0.001
	Differential-Balanced	28.60 [20.94, 36.27]	0.004
	Shared-Balanced	−2.47 [−7.57, 2.64]	1.000
S2 left − right	Differential-Fixed	49.63 [36.64, 62.62]	0.004
	Shared-Fixed	42.95 [28.17, 57.72]	0.009
	Differential-Balanced	26.41 [15.98, 36.84]	0.011
	Shared-Balanced	−0.34 [−9.16, 8.47]	1.000
P2 − P1	Differential-Fixed	79.62 [75.16, 84.09]	< 0.001
	Shared-Fixed	71.18 [63.80, 78.55]	< 0.001
	Differential-Balanced	35.69 [25.81, 45.57]	0.004
	Shared-Balanced	−0.83 [−5.27, 3.61]	1.000

Table 10 shows that Differential-Fixed, Shared-Fixed, and Differential-Balanced exhibited significant directional hand-role asymmetries across all four contrasts after Holm correction. The largest asymmetries were observed in the Fixed conditions, particularly Differential-Fixed, indicating strong dependence on the canonical right-tracking/left-holding allocation. In contrast, none of the Shared-Balanced contrasts was significant, and all four differences were close to zero. This pattern is consistent with substantially greater hand-role flexibility under Shared-Balanced training.