Toward Mobile Manipulation with Canal-Surface Shared Autonomy

Shalutha Rajapakshe
Idiap Research Institute and EPFL
Email: shalutha.rajapakshe@idiap.ch

Emmanuel Senft
Idiap Research Institute
Email: emmanuel.senft@idiap.ch

***Abstract*—Mobile manipulators have the potential to support people with disabilities by assisting with everyday household manipulation tasks. However, full autonomy in homes remains difficult because domestic environments are cluttered, dynamic, and shaped by user-specific preferences. Shared autonomy offers a practical alternative, allowing the robot to perform structured parts of a task while the user provides guidance when local context or preference matters. We present early work toward extending canal-surface shared autonomy to mobile manipulators, building on our previous framework, GeoSACS. As a first step, we use a simple and interpretable design in which the demonstrated canal remains fixed, while the mobile base and arm share the motion needed to follow it. This is achieved through a deterministic base-arm decomposition, which preserves a stable relationship between joystick corrections and task-space motion. This user-centered design prioritizes interpretability and predictable base motion, while leaving full whole-body control as a future extension. More broadly, this work opens a path toward treating canal surfaces not only as local correction structures, but also as reusable task-level primitives for future assistive mobile manipulation systems.**

I. Introduction

Mobile manipulators are increasingly relevant for assistive robotics because they can operate across real-world workspaces, including homes, rather than being limited to fixed tabletop settings. Recent work has shown rapid progress in open-vocabulary mobile manipulation [25], household skill learning on mobile manipulators from real-home demonstrations [21], whole-body mobile manipulation through teleoperation and imitation learning [8], and assistive mobile manipulation with accessible shared-control interfaces[18]. Together, these developments position mobile manipulators as a promising foundation for user-centered household assistance.

However, homes remain difficult environments for full autonomy: they are cluttered, dynamic, and shaped by user-specific routines, preferences, and accessibility needs. Assistive robotics therefore cannot focus only on autonomous task success; robots must also preserve user agency, perceived control, comfort, and safety [7, 17].

Shared autonomy (SA) offers a practical alternative, allowing the robot to handle structured or physically demanding parts of a task while the user remains involved when local context, ambiguity, or preference matters [13, 9]. Designing SA then raises a key technical question: how should user input be mapped into robot motion? Assistive interfaces are often low-dimensional, such as joysticks, switches, head-worn sensors, or web-based controls [18, 16, 20]. The robot, however, must generate high-dimensional end-effector motion. This question becomes especially important for mobile manipulators, where the same end-effector motion can be produced by different combinations of base and arm movement. A user-centered controller must therefore not only solve the kinematics, but also maintain a mapping between user input and robot behavior that is understandable, predictable, and easy to correct.

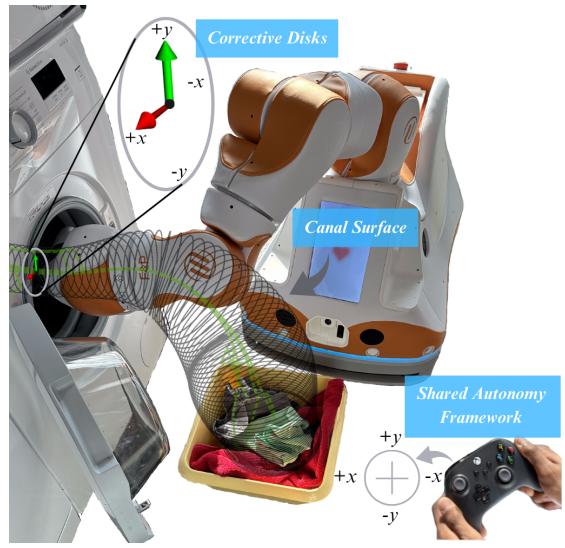

Fig. 1. Original GeoSACS canal-surface SA framework: the robot advances along a demonstrated canal while the user provides joystick corrections on local cross-sections. We build on this representation for mobile manipulators with coordinated base–arm motion.

GeoSACS addresses this dimensionality gap by representing demonstrated manipulation skills as canal surfaces [20]. A canal captures the nominal task progression as a centerline and defines local cross-sections where user corrections can be applied [10, 1]. The robot advances along the canal while the user provides 2D joystick corrections on the current cross-section, creating an interpretable correction space that preserves the demonstrated task structure. However, GeoSACS did not explicitly address how canal following should be coordinated with mobile-base motion.

This paper presents early work toward extending canal-surface SA to mobile manipulators. The central technical question is: if the canal is a task-space structure, how should it

be coordinated with a moving mobile base? Full whole-body control is a natural long-term direction, but immediately introducing whole-body optimization may obscure the user-facing correction structure that makes canal-based SA valuable. In our first implementation, informed by pilot insights, we make a deliberate user-centered design choice: the canal remains fixed in a demonstration frame, and desired end-effector motion is decomposed into a deterministic base component and an arm residual. This keeps the base contribution explicit and preserves a stable relationship between joystick corrections and visible robot behavior.

This conservative first step also points toward a broader direction for canal-surface SA. Many household tasks share similar geometric structure across different locations, such as moving objects between surfaces, loading laundry, placing dishes, or opening drawers. If canals can be coordinated with mobile-base motion, future systems could instantiate learned canal structures near different task locations through navigation and perception, while still allowing users to provide local corrections when object positions, target locations, or preferences differ from the original demonstration.

With that vision, our contributions are: (i) an initial formulation for extending canal-surface shared autonomy to mobile manipulators through a fixed task representation and deterministic base–arm decomposition; (ii) a Stretch 3 implementation that preserves both autonomous canal following and user-guided correction; and (iii) a discussion of how this first step can lead toward movable canals, navigation-conditioned reuse, and whole-body canal-following control.

II. Related Work

A. Assistive and user-centered mobile manipulation

Recent mobile-manipulation systems show that robots are increasingly able to operate beyond fixed tabletop settings. HomeRobot formalizes open-vocabulary mobile manipulation by combining navigation, perception, and manipulation in unseen environments [25]. Dobb-E demonstrates household manipulation skills on a Stretch robot from short demonstrations collected in real homes [21], while Mobile ALOHA highlights the value of coordinated base-arm motion for whole-body mobile manipulation [8]. Language-conditioned systems deployed on mobile robots, such as SayCan [2], PaLM-E [6], RT-2 [26], and TidyBot [24], further point toward reusable skills, semantic generalization, and personalization in household environments.

For assistive robotics, however, mobile manipulation is not only about autonomy or task success. Household assistance must also account for safety, comfort, user agency, and individual preferences. Prior work on activities of daily living (ADLs), such as feeding highlights the importance of human state and preference [7, 17], while HAT demonstrated that accessible teleoperation and shared-control assistance can support in-home mobile manipulation [18]. These works show that mobile manipulators can expand the physical reach of assistive robots, but users still need predictable ways to guide, correct, and personalize robot behavior. Our work builds on this need by extending canal-surface SA to mobile manipulators, combining the workspace benefits of a mobile base with an interpretable task-space correction structure.

B. Shared autonomy and low-dimensional input

SA is useful because real-world tasks often involve uncertainty, variability, and user preferences that are difficult to resolve autonomously. Hindsight optimization assists users under goal uncertainty [13], corrective SA lets users modify key robot state variables during autonomous behavior [9], and learned latent actions map low-dimensional joystick input into higher-dimensional robot action spaces [16]. These methods reduce user effort, but also raise a central design question: what should the user be allowed to correct, and how should simple input map to robot motion?

GeoSACS answers this mapping question geometrically using canal surfaces [20]. A task is encoded by a directrix and local cross-sections: the robot advances along the directrix, while the user corrects within the current cross-section. This preserves a meaningful relationship between 2D joystick input and task-space motion. For mobile manipulators, however, the mapping becomes more complex because the robot must also decide how the base contributes to end-effector motion. Broader human-robot interaction (HRI) work shows that users adapt when shared-control assistance changes robot behavior [4], that intervention depends on how clearly assistance is communicated [11], and that robot motion should be predictable or legible [5]. This supports our initial design choice: prioritizing an interpretable correction interface and deterministic base behavior before introducing full whole-body flexibility.

C. Base-arm coordination for canal-surface mobile manipulation

Extending canal-surface SA to mobile manipulators requires deciding how the mobile base and arm should share motion while preserving a predictable user correction interface. Whole-body control is a natural formulation because the base and arm jointly determine end-effector motion, and prior work has addressed this through redundancy resolution, task weighting, and whole-body inverse kinematics [22, 23, 14]. These methods are powerful and are an important direction for future canal-surface SA, but the resulting motion can depend on costs, weights, constraints, and the instantaneous robot configuration. In SA, this may make the mapping from user input to robot response harder to predict.

Our current controller therefore starts with deterministic base-arm decomposition, leaving full whole-body canal-following control as a future extension. This makes the mobile base's role explicit: the base contributes a predictable component of the end-effector motion, while the arm handles the remaining motion needed to follow the canal and user corrections. Once this coordination is established, canal-surface SA can be extended toward richer mobile manipulation settings, where navigation brings the robot near a task location, the canal is instantiated for the current task, and the base helps

both approach the canal and extend the reachable workspace during execution.

This direction also connects to learning from demonstration, where movement primitives and probabilistic skill representations encode demonstrated motions as reusable structures rather than only replayed as single trajectories [3, 12, 19]. Similarly, canal surfaces could eventually be reused across household tasks with similar geometric structure. For example, floor-to-table object transfer can share a common task pattern even when the object, location, table, or obstacles differ. In future work, movable or relaxed canals could adapt such structures across locations while preserving a clear correction interface when the generalized motion is insufficient.

III. Method: Extending Canal-Surface SA to Mobile Manipulators

A. Canal representation in a fixed demonstration frame

We represent the demonstrated skill as a canal surface following GeoSACS [20]. Let s index progress along the canal, $c(s) \in \mathbb{R}^3$ be the directrix, $r(s)$ be the cross-section radius, and x_s, y_s be the correction axes. A user-corrected point on the current cross-section is

$$p_C(s, \alpha, \beta) = c(s) + r(s)(\alpha x_s + \beta y_s), \quad \alpha^2 + \beta^2 \leq 1. \quad (1)$$

The robot advances along s, while the user controls (α, β) through a low-dimensional input device.

In this first mobile-manipulator extension, the canal remains fixed in a demonstration frame. The robot base moves relative to this frame, and odometry relates the current base pose to the demonstrated task structure. This lets us focus on the first coordination problem: how the base and arm should share motion while preserving an interpretable correction interface.

B. Deterministic base-arm decomposition

For a mobile manipulator, the end-effector motion can be produced by both the mobile base and the arm:

$$\dot{x}_{ee} = J_b(q)\dot{q}_b + J_a(q)\dot{q}_a. \quad (2)$$

This redundancy is useful because it allows the robot to distribute motion across the whole body. However, in an SA setting, it can also make the robot response harder to predict: depending on the controller weights, constraints, and current configuration, the same user correction may result in different combinations of base motion, arm motion, or both.

To keep the first extension interpretable, we resolve this allocation explicitly. Let Δp be the desired end-effector displacement in the fixed demonstration frame, and let e_b be the selected base translation direction expressed in the same frame. We decompose the desired displacement into a base component and an arm residual:

$$\Delta p_b = (e_b^T \Delta p) e_b, \qquad \Delta p_a = \Delta p - \Delta p_b. \quad (3)$$

The mobile base is commanded only along Δp_b, while the arm executes the remaining motion Δp_a. This makes the role of the base deterministic and easy to explain: the base handles the selected translational part of the task-space motion, and the arm handles the rest.

This design is not intended to replace whole-body control. Rather, it is a conservative first step that preserves a stable relationship between joystick corrections and visible robot behavior. Once this base–arm coordination is better understood and evaluated with users, the formulation can be relaxed toward whole-body canal-following control.

C. Execution modes, user initiative, and correction

The original GeoSACS formulation separates robot-led canal advancement from user correction on the current cross-section. We preserve this structure for mobile manipulation by enforcing explicit motion ownership between autonomous canal following and user-guided correction. In autonomous mode, the robot advances along the canal using the deterministic base-arm split. In correction mode, joystick input updates the current cross-section point and pauses autonomous advancement until the user resumes. This prevents autonomous and user-guided commands from competing, which is important when both the base and arm can contribute to motion.

Motivated by pilot feedback, we also implemented a user-led variant that shifts more initiative to the user. In the robot-led version, the robot advances automatically and the user corrects only when needed. In the user-led version, the canal still provides the task-space structure, but the user controls progression using hold-to-run inputs; when no progression command is active, the robot holds its state. This keeps the same canal representation and base-arm decomposition, but changes who controls movement along the canal. We treat it as an implementation option rather than a final design choice, pending future user studies.

Finally, we extend the original GeoSACS interaction with a lightweight orientation correction for tasks where gripper roll matters. In GeoSACS, orientation is primarily inherited from the demonstrated trajectory. Here, we allow the user to apply a persistent roll offset on top of the demonstrated orientation. Let $q_G(s)$ be the demonstrated GeoSACS orientation and ρ be a user-controlled roll offset around the local end-effector axis. The commanded orientation is:

$$q_{cmd}(s) = q_G(s) \otimes q_x(\rho), \quad (4)$$

where $q_x(\rho)$ is the local roll quaternion. The offset persists across cross-sections until the user changes it. This preserves the demonstrated orientation as the default, while allowing the user to adapt the gripper roll to the task at hand without redefining the full trajectory.

IV. Current Status and Pilot Insights

We implemented the proposed canal-surface SA extension on the Stretch 3 mobile manipulator [15]. The current prototype supports robot-led canal following, user-guided cross-section correction, deterministic base-arm decomposition, persistent roll-offset correction, and a user-led traversal variant. It is currently used as a formative prototype rather than a completed evaluated system.

Early pilot sessions highlighted differences in preferred user initiative. Some participants preferred robot-led traversal for simple pick-and-place motions, where the task structure was clear and they were comfortable letting the robot handle the main motion. Others preferred user-led traversal through hold-to-run gamepad inputs, especially, for contact-sensitive tasks, such as approaching and pulling a drawer handle, where finer control was desired near the interaction region. We treat both the robot-led and user-led initiates as part of the design space for future evaluation rather than as final design choices.

Pilot feedback also supported our decision to make mobile-base motion explicit and predictable. When the base moved while participants controlled the end-effector, three out of four pilot participants mentioned that base motion could be confusing. Participants appeared to better understand the behavior when the base contribution was aligned with the base motion direction, rather than used as an implicit compensation mechanism. This supports deterministic base-arm decomposition as an initial user-centered design choice, although further studies are needed to evaluate systematically.

Pilot feedback also shaped the orientation correction design. In an earlier implemented version, user-applied roll corrections were gradually returned to the demonstrated default orientation as the robot moved back toward the middle of the canal after picking or placing. Participants found this confusing because the gripper changed without direct input. We therefore store roll correction as a persistent offset on top of the demonstrated orientation, as described in Section III-C, making user-driven orientation changes predictable and reversible.

Together, these pilot insights suggest that extending canal-surface SA to mobile manipulators should preserve an understandable relationship between user input and robot behavior. Future studies will examine how traversal mode, task type, and deterministic base-arm coordination affect perceived predictability and control.

V. Discussion and Next Steps

The main contribution of this work is not the specific one-dimensional base split alone, but the first step it provides toward canal-surface SA for mobile manipulation. The current implementation keeps the canal fixed and introduces mobile-base motion in a predictable way, so that the user-facing correction interface remains interpretable. This choice is intentionally conservative: before moving toward richer base motion or full whole-body control, we first establish a simple coordination strategy between the canal, the arm, and the mobile base.

This foundation opens a path toward navigation-conditioned canal reuse. In future systems, a canal could be represented in a task frame T and instantiated in the world frame W using a task transform:

$$p_W(s, \alpha, \beta) = R_T^W p_T(s, \alpha, \beta) + t_T^W . \tag{5}$$

This would allow a learned canal structure, such as a floor-to-table, table-to-table, loading, wiping, or drawer-interaction motion, to be reused across different household locations. Navigation could bring the robot near the task, perception could estimate the relevant task frame, and the canal could provide the manipulation structure. The mobile base would then help position the robot with respect to the canal and extend the reachable workspace, while the user would still provide local corrections when object positions, target locations, obstacles, or preferences differ from the original demonstration.

This paper does not solve movable canal reuse. Instead, it identifies and implements an earlier step needed for that direction: coordinating a fixed canal with mobile-base motion while preserving an interpretable correction interface. The staged technical path is therefore: first, fixed canals with deterministic base-arm decomposition; second, relaxed canal constraints that allow the robot to operate beyond the original canal boundary when useful; third, movable task-frame canals that can be instantiated after navigation and perception; and fourth, whole-body canal-following control that optimizes base and arm motion while preserving the user-facing correction structure.

This staged path also makes evaluation clearer. Future studies can compare robot-led and user-led traversal, deterministic base-arm decomposition, richer base motion, and full whole-body control. These comparisons should consider not only task success and completion time, but also workload, perceived control, predictability, intervention timing, and correction behavior. Such evidence is needed before replacing the current deterministic controller with a more expressive but potentially less transparent one.

VI. Conclusion

We presented early work toward mobile manipulation with canal-surface SA. The current controller keeps the canal fixed in a demonstration frame, relates the mobile base to this frame, and applies a deterministic base-arm decomposition to preserve predictable joystick corrections. We also implemented robot-led and user-led traversal options, as well as a persistent roll-offset correction for task-specific gripper adjustment. Pilot insights suggest that predictable base motion, explicit user initiative, and persistent user-controlled orientation changes are important design considerations for extending canal-surface SA to mobile manipulators.

This work is a first step rather than an endpoint. It provides a technical foundation for future movable canals, navigation-conditioned canal reuse, relaxed operating spaces, and whole-body canal-following control. More broadly, we argue that canal surfaces can become useful task-space structures for assistive mobile manipulators, supporting household manipulation while preserving the user agency and interpretability needed for real-world use.

References

[1] S. Reza Ahmadzadeh, Roshni Kaushik, and Sonia Chernova. Trajectory learning from demonstration with canal surfaces: A parameter-free approach. In *2016 IEEE-RAS 16th International Conference on Humanoid Robots*

(Humanoids), 2016. doi: 10.1109/HUMANOIDS.2016.7803328.

[2] Michael Ahn et al. Do As I Can, Not As I Say: Grounding Language in Robotic Affordances. In *Proceedings of the 6th Conference on Robot Learning*, 2022. URL https://say-can.github.io/.

[3] Brenna D. Argall, Sonia Chernova, Manuela Veloso, and Brett Browning. A Survey of Robot Learning from Demonstration. *Robotics and Autonomous Systems*, 57 (5):469–483, 2009. doi: 10.1016/j.robot.2008.10.024. URL https://doi.org/10.1016/j.robot.2008.10.024.

[4] Reuben M. Aronson and Elaine Schaertl Short. Intentional User Adaptation to Shared Control Assistance. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pages 4–12. ACM, 2024. doi: 10.1145/3610977.3634953. URL https://doi.org/10.1145/3610977.3634953.

[5] Anca D. Dragan, Kenton C. T. Lee, and Siddhartha S. Srinivasa. Legibility and Predictability of Robot Motion. In *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction*, pages 301–308. IEEE, 2013. doi: 10.1109/HRI.2013.6483603. URL https://doi.org/10.1109/HRI.2013.6483603.

[6] Danny Driess et al. PaLM-E: An Embodied Multimodal Language Model. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 8469–8488. PMLR, 2023. URL https://proceedings.mlr.press/v202/driess23a.html.

[7] Zackory Erickson, Vamsee Gangaram, Ariel Kapusta, C. Karen Liu, and Charles C. Kemp. Assistive Gym: A Physics Simulation Framework for Assistive Robotics. In *2020 IEEE International Conference on Robotics and Automation*, pages 10169–10176. IEEE, 2020. doi: 10.1109/ICRA40945.2020.9197411. URL https://doi.org/10.1109/ICRA40945.2020.9197411.

[8] Zipeng Fu, Tony Z. Zhao, and Chelsea Finn. Mobile ALOHA: Learning Bimanual Mobile Manipulation using Low-Cost Whole-Body Teleoperation. In *Proceedings of The 8th Conference on Robot Learning*, volume 270 of *Proceedings of Machine Learning Research*, pages 4066–4083. PMLR, 2025. URL https://proceedings.mlr.press/v270/fu25b.html.

[9] Michael Hagenow, Emmanuel Senft, Robert Radwin, Michael Gleicher, Bilge Mutlu, and Michael Zinn. Corrective Shared Autonomy for Addressing Task Variability. *IEEE Robotics and Automation Letters*, 6(2):3720–3727, 2021. doi: 10.1109/LRA.2021.3064500. URL https://doi.org/10.1109/LRA.2021.3064500.

[10] D. Hilbert and S. Cohn-Vossen. Geometry and the Imagination. *Physics Today*, 6(5):19–19, 05 1953. ISSN 0031-9228.

[11] Joshua Hoegerman, Shahabedin Sagheb, Benjamin A. Christie, and Dylan P. Losey. Aligning Learning with Communication in Shared Autonomy. In *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 11530–11536. IEEE, 2024. doi: 10.1109/IROS58592.2024.10801897. URL https://doi.org/10.1109/IROS58592.2024.10801897.

[12] Auke Jan Ijspeert, Jun Nakanishi, Heiko Hoffmann, Peter Pastor, and Stefan Schaal. Dynamical Movement Primitives: Learning Attractor Models for Motor Behaviors. *Neural Computation*, 25(2):328–373, 2013. doi: 10.1162/NECO_a_00393. URL https://doi.org/10.1162/NECO_a_00393.

[13] Shervin Javdani, Henny Admoni, Stefania Pellegrinelli, Siddhartha S. Srinivasa, and J. Andrew Bagnell. Shared Autonomy via Hindsight Optimization for Teleoperation and Teaming. *The International Journal of Robotics Research*, 37(7):717–742, 2018. doi: 10.1177/0278364918776060. URL https://doi.org/10.1177/0278364918776060.

[14] Akihiro Kanazawa, Shuichi Akizuki, Kenji Hashimoto, Wataru Takahashi, Tomohito Tanaka, and Takamitsu Moriyama. Whole-Body Inverse Kinematics Robust to Base Position Control Error in Mobile Manipulators. *Advanced Robotics*, 2024. doi: 10.1080/01691864.2024.2369816. URL https://doi.org/10.1080/01691864.2024.2369816.

[15] Charles C. Kemp, Aaron Edsinger, Henry M. Clever, and Blaine Matulevich. The design of stretch: A compact, lightweight mobile manipulator for indoor human environments. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 3150–3157, 2022. doi: 10.1109/ICRA46639.2022.9811922.

[16] Dylan P. Losey, Krishnan Srinivasan, Ajay Mandlekar, Animesh Garg, and Dorsa Sadigh. Controlling Assistive Robots with Learned Latent Actions. In *2020 IEEE International Conference on Robotics and Automation*, pages 378–384. IEEE, 2020. doi: 10.1109/ICRA40945.2020.9197197. URL https://doi.org/10.1109/ICRA40945.2020.9197197.

[17] Benjamin A. Newman, Reuben M. Aronson, Siddhartha S. Srinivasa, Kris Kitani, and Henny Admoni. HARMONIC: A Multimodal Dataset of Assistive Human-Robot Collaboration. *The International Journal of Robotics Research*, 41(1):3–11, 2022. doi: 10.1177/02783649211050677. URL https://doi.org/10.1177/02783649211050677.

[18] Akhil Padmanabha, Janavi Gupta, Chen Chen, Jehan Yang, Vy Nguyen, Douglas J. Weber, Carmel Majidi, and Zackory Erickson. Independence in the Home: A Wearable Interface for a Person with Quadriplegia to Teleoperate a Mobile Manipulator. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pages 542–551. ACM, 2024. doi: 10.1145/3610977.3634964. URL https://doi.org/10.1145/3610977.3634964.

[19] Alexandros Paraschos, Christian Daniel, Jan Peters, and Gerhard Neumann. Probabilistic Movement Primitives. In *Advances in Neural Information Processing Systems*, volume 26, 2013. URL

https://proceedings.neurips.cc/paper/2013/hash/e53a0a2978c28872a4505bdb51db06dc-Abstract.html.

[20] Shalutha Rajapakshe, Atharva Dastenavar, Michael Hagenow, Jean-Marc Odobez, and Emmanuel Senft. GeoSACS: Geometric Shared Autonomy via Canal Surfaces. In *Proceedings of the 21st ACM/IEEE International Conference on Human-Robot Interaction*, pages 744–753. ACM, 2026. doi: 10.1145/3757279.3785579. URL https://doi.org/10.1145/3757279.3785579.

[21] Nur Muhammad Mahi Shafiullah, Anant Rai, Haritheja Etukuru, Yiqian Liu, Ishan Misra, Soumith Chintala, and Lerrel Pinto. On Bringing Robots Home. *arXiv preprint arXiv:2311.16098*, 2023. URL https://arxiv.org/abs/2311.16098.

[22] Frederico Fernandes Afonso Silva and Bruno Vilhena Adorno. Whole-Body Control of a Mobile Manipulator Using Feedback Linearization and Dual Quaternion Algebra. *Journal of Intelligent & Robotic Systems*, 91(2): 249–262, 2018. doi: 10.1007/s10846-017-0686-3. URL https://doi.org/10.1007/s10846-017-0686-3.

[23] Zhangjie Tu, Tianwei Zhang, Lei Yan, and Tin Lun Lam. Whole-Body Control for Velocity-Controlled Mobile Collaborative Robots Using Coupling Dynamic Movement Primitives. *arXiv preprint arXiv:2203.03296*, 2022. URL https://arxiv.org/abs/2203.03296.

[24] Jimmy Wu, Rika Antonova, Adam Kan, Marion Lepert, Andy Zeng, Shuran Song, Jeannette Bohg, Szymon Rusinkiewicz, and Thomas Funkhouser. TidyBot: Personalized Robot Assistance with Large Language Models. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2023. URL https://tidybot.cs.princeton.edu/.

[25] Sriram Yenamandra et al. HomeRobot: Open-Vocabulary Mobile Manipulation. In *Proceedings of The 7th Conference on Robot Learning*, volume 229 of *Proceedings of Machine Learning Research*, pages 1975–2011. PMLR, 2023. URL https://proceedings.mlr.press/v229/yenamandra23a.html.

[26] Brianna Zitkovich et al. RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control. In *Proceedings of The 7th Conference on Robot Learning*, volume 229 of *Proceedings of Machine Learning Research*, pages 2165–2183. PMLR, 2023. URL https://proceedings.mlr.press/v229/zitkovich23a.html.