

Learning Quiet Walking for a Small Home Robot

Ryo Watanabe^{1,2}, Takahiro Miki¹, Fan Shi^{1,3}, Yuki Kadokawa^{1,4}, Filip Bjelonic¹,
Kento Kawaharazuka^{1,5}, Andrei Cramariuc¹ and Marco Hutter¹

Abstract—As home robotics gains traction, robots are increasingly integrated into households, offering companionship and assistance. Quadruped robots, particularly those resembling dogs, have emerged as popular alternatives for traditional pets. However, user feedback highlights concerns about the noise these robots generate during walking at home, particularly the loud footstep sound. To address this issue, we propose a sim-to-real based reinforcement learning (RL) approach to minimize the foot contact velocity highly related to the footstep sound. Our framework incorporates three key elements: learning varying PD gains to actively dampen and stiffen each joint, utilizing foot contact sensors, and employing curriculum learning to gradually enforce penalties on foot contact velocity. Experiments demonstrate that our learned policy achieves superior quietness compared to a RL baseline and the carefully handcrafted Sony commercial controllers. Furthermore, the trade-off between robustness and quietness is shown. This research contributes to developing quieter and more user-friendly robotic companions in home environments.

I. INTRODUCTION

Home robotics is a rapidly growing trend driven by advancements in actuators, sensors, and artificial intelligence. These developments have allowed home robots to perform a wider range of tasks and interact with humans in more meaningful ways [1]–[4]. For example, aibo [1], a robotic pet introduced by Sony, is capable of learning and developing its own unique personality through interactions with its home environment and owner. This emotional connection with robotic pets demonstrates the potential for robots to play a significant role in our daily lives [5]–[8].

Reinforcement learning (RL) has recently been applied to robotic locomotion, significantly improving robustness in a variety of terrains [10]–[15]. While these advances have proven beneficial for outdoor use, bringing similar robustness and adaptability to home robots remains an important goal. However, much of the research in RL based legged locomotion has prioritized robustness and efficiency, with little focus on reducing the footstep noise generated during walking. In home environments, where quiet operation is critical, reducing walking noise is essential. For example, one of the main concerns of aibo users is that the walking sound is too loud. This feedback highlights the need for quieter locomotion policies in home robots.

¹Ryo Watanabe, Takahiro Miki, Fan Shi, Yuki Kadokawa, Filip Bjelonic, Kento Kawaharazuka, Andrei Cramariuc, and Marco Hutter are with the Robotic Systems Lab, Department of Mechanical Engineering, ETH Zurich, Switzerland.

²Ryo Watanabe is at Sony Group Corporation, Japan

³Fan Shi is at National University of Singapore

⁴Yuki Kadokawa is at Nara Institute of Science and Technology, Japan

⁵Kento Kawaharazuka is at The University of Tokyo, Japan

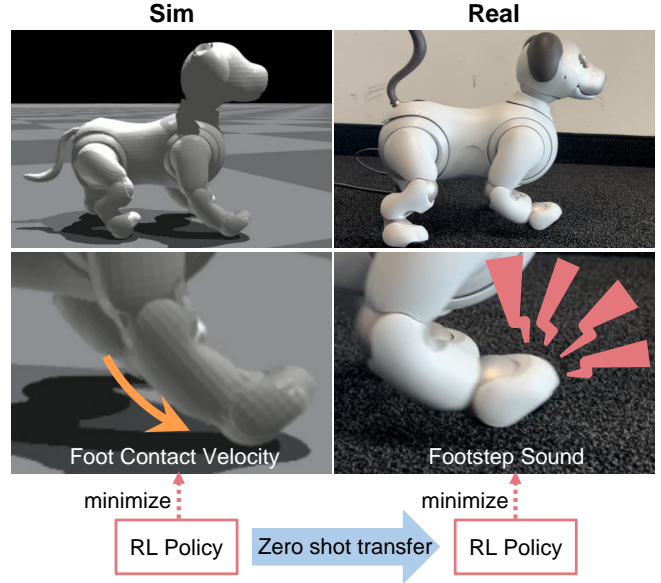


Fig. 1. For the aibo small home robot pictured above, we design a sim-to-real based RL approach to minimize the foot contact velocity in the physics simulator, which highly correlates with the footstep sound in the real world to achieve quiet walking. Project webpage: <https://sony.github.io/QuietWalking/>

In this work, we focus on minimizing footstep noise by using sim-to-real based RL to reduce foot contact velocity, which serves as a proxy for footstep sound reduction, as shown in Fig. 1. To achieve this, we introduce three key elements aimed at effectively reducing foot contact velocity during locomotion. First, we allow the RL network to dynamically adjust the PD gains of the actuators, allowing for better joint damping and stiffening. Second, we incorporate foot switch contact sensors that detect when the robot’s foot touches the ground to selectively dampen joints during the contact phase and stiffen them during the stance phase for body support. Third, we employ a curriculum learning approach that initially trains the robot on basic locomotion and progressively applies penalties to noisy walking such as foot contact velocity to ensure quieter walking.

We validate our approach through hardware experiments with Sony’s quadruped robot aibo. In these experiments, we show that foot contact velocity correlates with footstep noise, and our method produces the quietest locomotion compared to both Sony’s carefully hand-tuned controllers and other RL baselines. In addition, we explore the trade-off between robustness and quietness in locomotion and show that this balance can be adjusted using Domain Randomization (DR) parameters.

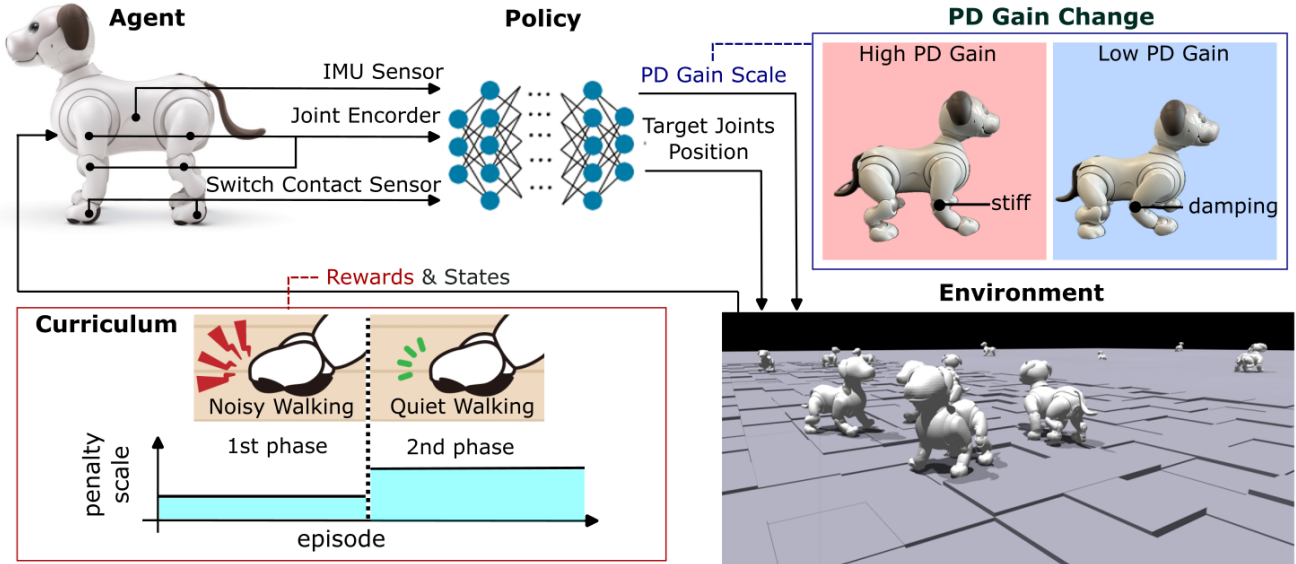


Fig. 2. System overview of the RL training framework for learning quiet locomotion with aibo. The agent, aibo, uses its IMU, joint encoders, and switch contact sensors to compute observations. As an action for 12 joints, the policy outputs target joint position and the joint PD gain scale, which enables the tracking of the target joints with high PD gain and the damping of the joints with low PD gain. On the right is the Isaac Gym [9] simulation environment, which we leverage for parallel training on GPUs of multiple agents. The reward scales are divided into two phases, where aibo first learns just to walk, and then in the second phase, it adapts its walk to be quieter.

In summary, our main contributions are as follows:

- Demonstrate superior quiet locomotion in a small home robot by minimizing physical parameters, such as foot contact velocity, that correlate with footstep sound using a sim-to-real RL approach.
- Develop an RL framework that integrates three core elements: adaptive PD gains, foot contact sensors, and curriculum learning, resulting in an effective policy for quiet walking.
- Identify and demonstrate the trade-off between robustness and quietness in robot locomotion, with DR parameters providing a way to balance these competing objectives.

II. RELATED WORK

A. Reinforcement Learning Locomotion

RL has been successfully used to generate robust locomotion policies for quadrupedal robots [10]–[13]. All of the previously mentioned works focus on robustness but ignore the user-interaction aspect of the locomotion policy, namely the amount of noise the robot generates when moving around. Even locomotion policies designed to be energy efficient do not inadvertently reduce the amount of noise they produce, as seen in the work of Yang *et al.* [16]. Other works on energy efficient gaits [17] or imitation learning-based gaits do not address or evaluate the impact noise of the learned policy [18], [19]. While many of these works show promising results for robustness and efficiency, they are not designed to suppress loudness, which we aim to solve.

B. Sound Engineering in Robotics

Sound plays a crucial role in how humans perceive and interact with various products. Extensive engineering

efforts have been made in the automotive industry [20] to optimize the acoustic experience for drivers. Recently, the robotics industry has also recognized the importance of sound engineering, with research focusing on robotic arms [21], [22], bipedal robots [23], rolling robots [24], and servomotors [25]. These studies have investigated how the generated sounds influence human comfort and perception during human-robot interactions, leading to the development of more pleasant and even cute motion patterns [26]. However, previous works have not thoroughly explored the acoustic impact of quadruped robot footsteps. This paper aims to address this gap by focusing on evaluating and reducing the sound generated by the footsteps of quadruped robots.

III. METHOD

To minimize foot contact velocity in the physics simulator correlated to the footstep sound in the real, we propose a sim-to-real deep RL framework which synergistically incorporates three key components as shown in Fig. 2.

The training network architecture consists of three fully connected layers, each with 128 hidden units and Exponential Linear Unit (ELU) activation functions, and utilizes the Proximal Policy Optimization (PPO) [27]. Both the actor and critic networks are implemented as Multi-Layer Perceptrons (MLPs), adhering to the architecture and PPO hyperparameters established in prior work [15]. The simulation and control loop frequencies are set to 400 Hz and 100 Hz, respectively, using the Isaac Gym framework [9].

A. Actions

As shown in Fig. 2, the policy outputs two types of actions for the 12 leg joints: target joint positions and PD gain scales. a^* and $gain^*$ represent target joint positions and

TABLE I
POLICY OBSERVATION

Entry	Symbol	Noise level	Units	Dimensions
Joint positions	φ	0.01	rad	12
Joint velocities	$\dot{\varphi}$	1.5	rad/s	12
Last action (joint position)	\mathbf{a}^*	0.0	rad	12
Last action (PD gain scale)	\mathbf{gain}^*	0.0	—	12
Foot contact state	\mathbf{f}_c	0.0	—	4
Gravity orientation	\mathbf{g}	0.05	m/s^2	3

PD gain scales, respectively. We chose target joint position and PD gain scale as actions instead of directly estimating torque for two primary reasons. First, previous work [28] reported that torque control as an action wasn't efficient for learning locomotion. Secondly, previous work [29] found that estimating target joint positions and PD gains is more suitable for sensitive contact tasks that require soft ground contact.

In our approach, the proportional P_i and derivative D_i gains are calculated using the following formulas:

$$\begin{aligned} P_i &= P^* + \alpha \sigma(x_i) \\ D_i &= D^* + \beta \sigma(x_i) \end{aligned} \quad (1)$$

where σ is the sigmoid function, α and β are scaling factors, P^* and D^* are the nominal gains, and i represents each leg joint. The specific parameter values are $\alpha = 4.0$, $\beta = 0.02$, $P^* = 3.0$, and $D^* = 0.03$. Our locomotion policy outputs only one gain scale factor x_i for both the proportional P_i and derivative D_i parts of the PD controller. While this approach of outputting a single gain scale from the policy is consistent with the previously mentioned work [29], this method is the first application to locomotion tasks in quadruped robots to explore the quietness.

B. Observations

Policy observations and noise levels are detailed in Table I. One of the key elements of our approach is utilizing the switch contact sensor to determine when to stiffen the joint from its soft state during the contact phase. Since aibo is a consumer-grade small robot designed to be affordable for the average household, it does not have force and torque sensors. Without these sensors, there is a traditional way to estimate external force to detect contact using model-based methods [30]. However, such methods are not sufficiently accurate to determine when a small robot contacts the ground, as it is difficult to precisely identify friction and develop an accurate robot model for commercially available mass production robots. Instead of estimating external force, we utilized switch contact sensors to detect the contact directly. Although the switch contact sensors provide binary output, they are still useful for the contact sensitive task such as quiet walking.

C. Rewards

Reward functions and their relative scales are described in Table II. For task rewards, \mathbf{v}_b , \mathbf{v}_b^* , \mathbf{w}_b , and \mathbf{w}_b^* denote the

measured and desired linear velocities in the base frame and the measured and desired angular velocities in the base frame, respectively. Penalty rewards help suppress inappropriate or jerky motions. φ , τ , \mathbf{v}_f , $n_{collision}$ and $\mathbf{t}_{f,air}$ represent joint positions, joint torques, foot velocities, number of collisions and airtime of foot, respectively.

We incorporate three additional terms related to the footstep sound for quiet walking. Biomechanics research indicates that footstep sound correlates with foot contact velocity [31]. When the foot contacts the ground, the kinetic energy, $\frac{1}{2}m\mathbf{v}_{f,xyz}^2$, is converted into various forms of energy, including sound energy which produces the footstep sound. While foot contact velocity is the dominant parameter related to footstep sound, we also include joint acceleration and base angular acceleration penalties to support this objective. We observe that sudden changes in joint angles and base orientation often follow loud step sounds, as noisy walking causes the robot to bump into the ground.

D. Curriculum Learning

Initial attempts at training without curriculum learning proved unsuccessful, resulting in either non-convergent training or the aibo learning to remain stationary, a strong local minimum that minimizes noise but fails to achieve the desired locomotion. To address these challenges, we implemented a curriculum learning approach [32]. Inspired by a recent study proposing the division of a training episode into multiple stages to tackle the explore vs exploit dilemma [33], [34], we divide the training episode into two distinct stages: noisy walking and quiet walking stage.

During the noisy walking phase, we applied lower weights to the noise-related penalties, as shown in Table II. This strategy allowed the robot to prioritize learning basic locomotion and velocity command tracking, even if significant footstep sounds are generated. The transition to the quiet walking phase is triggered when the agents achieve the sum of episodic reward about linear and angular velocity tracking is more than 1.5. In the quiet walking phase, we increased the penalties associated with foot contact velocity, base angular acceleration, and joint acceleration. Given that foot contact velocity is the most critical parameter for footstep sound reduction, its penalty was increased 5 times compared to the noisy walking phase, while other penalties were only doubled as shown in Table II. By combining reward shaping with this two-phase curriculum learning approach, we successfully enabled the robot to develop a quiet locomotion policy that minimizes the footstep sound while preserving its capacity to execute velocity commands effectively.

E. Domain Randomization

We employ DR [35] to enhance the sim-to-real transfer of the policy. This involves procedurally generating various parameters in the simulation environment, such as the height of the steps, the ground friction coefficient, and the robot base mass. Additionally, we utilize disturbances during the episodes by applying an impulse signal to the robot's base velocity every 4 seconds and random external forces and

TABLE II
REWARD FUNCTIONS AND SCALES AT EACH CURRICULUM LEARNING PHASE

Category	Reward	Definition	Scale in noisy walking phase	Scale in quiet walking phase
Task Rewards	Linear velocity tracking	$\exp(-\frac{1}{0.06} \ \mathbf{v}_{b,xy}^* - \mathbf{v}_{b,xy}\ ^2)$	1.0	1.0
	Angular velocity tracking	$\exp(-\frac{1}{0.06} \ w_{b,z}^* - w_{b,z}\ ^2)$	1.0	1.0
Penalty Rewards	Joint torque	$\ \boldsymbol{\tau}\ ^2$	-0.015	-0.015
	Base linear velocity z	$\ v_{b,z}\ ^2$	-3.0	-3.0
	Base orientation	$\ \boldsymbol{\theta}_{b,xy}\ ^2$	-5.0	-5.0
	Base angular velocity	$\ \dot{\boldsymbol{\theta}}_{b,xy}\ ^2$	-0.05	-0.05
	Foot slippage	$\ \mathbf{v}_{f,xy}\ ^2$	-0.15	-0.15
	Self-collisions	$n_{collision}$	-10.0	-10.0
	Foot air time	$\sum_i (t_{f,air} - 0.2)$	2.0	2.0
	Joint target difference	$\ \mathbf{a}_{t-1}^* - \mathbf{a}_t^*\ ^2$	-0.02	-0.02
	PD gain scale difference	$\ \mathbf{gain}_{t-1}^* - \mathbf{gain}_t^*\ ^2$	-0.005	-0.005
Noisy Walking Penalties	Foot contact velocity	$\ \mathbf{v}_{f,xyz}\ ^2$	-5.0	-25.0
	Joint acceleration	$\ \ddot{\boldsymbol{\phi}}\ ^2$	-2e-7	-4e-7
	Base angular acceleration	$\ \ddot{\boldsymbol{\theta}}_{b,xy}\ ^2$	-5e-5	-1e-4

TABLE III
RANDOMIZED SIMULATION PARAMETERS

Randomized Variables	unit	min	max
Robot base mass	kg	-0.4	0.4
Velocity disturbance	m/s	-0.2	0.2
External force	N	0.0	0.4
External torque	Nm	0.0	0.1
Terrain height	m	0.002	0.01
Friction coefficient	[-]	0.4	0.7

torques to the base. The ranges for the randomized parameters are shown in Table III.

The extent of randomization significantly influences the resulting policy’s characteristics. While more extensive DR typically leads to a more robust gait, it can potentially compromise the quietness of locomotion. Previous research on RL-based racing drones [36] has mentioned the trade-off between robustness and performance depending on the degree of DR. Similarly, for quiet walking, it is crucial to strike an optimal balance when determining the level of DR.

IV. EXPERIMENTS AND RESULTS

We evaluate the quietness by using the microphone in the robot depending on the measured base velocity. By showing that the selective reward such as the foot contact velocity, base angular acceleration and joint angular acceleration at the proposed policy are smaller than the baseline policy, we prove those observable parameters in a physics simulator correlate to the footstep sound in the real. As part of our ablation study, we demonstrate the trade-off between robustness and quietness by validating three key factors of our framework and the effectiveness of DR.

A. Quietness Evaluation

We evaluate the performance of our proposed quiet RL walking policy and compare it to three other methods: a base-

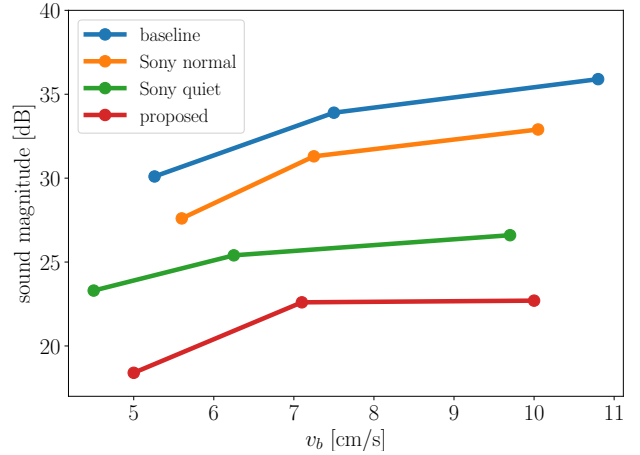


Fig. 3. Comparison of the average sound magnitude between 20 Hz and 20 kHz for the measured base velocity of aibo. The baseline and propose are RL policies without and with our three key factors for each. Sony normal and Sony quiet are the commercial controllers provided by Sony.

line RL policy and two commercial locomotion controllers provided by Sony. The baseline RL policy is trained using the standard rewards from the previous work [15] which we tuned to adapt to the small-scale robot, aibo. Unlike our proposed policy, the baseline RL policy does not output PD gain scale, does not include rewards for noisy walking penalties, and does not employ curriculum learning during training. Sony’s two baseline controllers are designed for normal walking and quiet walking, respectively. As these are proprietary algorithms, we compare our method to them without knowledge of their internal workings.

To evaluate the sound level of each locomotion policy, we use one of the microphones on the rear side of aibo’s head, which has a sampling frequency of 48 kHz. We focus our analysis on the standard human audible range between 20 Hz and 20 kHz [37], as this range is most relevant for assessing perceived noise levels. In this experiment, all locomotion controllers are set to track the same base velocity moving forward for the comparison. We preprocess the data using

TABLE IV
ANALYSIS OF NOISY WALKING PENALTIES AT SIMULATOR

Physical Term	Unit	Baseline	Proposed
Foot contact velocity	m/s	0.417	0.123
Joint acceleration	rad/s ²	114.3	76.7
Base angular acceleration	rad/s ²	57.2	23.7

a Hamming window and then use Welch’s method [38] to perform the frequency analysis with Fast Fourier Transforms (FFTs). We use a window size of 4096 samples corresponding to approximately 85 milliseconds. For the power analysis, we use as a reference a 10^{-10} amplitude sinusoid for each frequency bin.

The sound level of walking increases with higher commanded base velocities, as the robot has less time to slow its feet before contact while maintaining the commanded velocity. Therefore, we experimented to measure the average sound level for each policy at different measured linear velocity commands. It is important to note that the microphone is approximately 10 cm away from the feet, whereas the sound level would be much lower for a human standing further away. We only use the human audible frequency range and a 1 kHz sinusoidal reference signal with an amplitude of 1.0. Figure 3 shows that the quietness of proposed method outperforms the other methods such as the baselines and Sony commercial controllers at different measured velocities.

B. Analysis of Noisy Walking Penalties Rewards

To validate our approach of minimizing noisy walking penalties in the simulator for reducing aibo’s footstep sound in the real world, as illustrated in Fig. 1, we compared the noisy walking penalties between the baseline and proposed policies. Table IV presents the average values for 10 seconds calculated using the equations in Table II in the Isaac Gym simulator. The policies used for the baseline and proposed are those employed in Fig. 3. As shown in Table IV, our analysis reveals that the proposed policy, which produces quieter walking compared to the baseline, exhibits lower values across all observed parameters. As mentioned in subsection IV-A, the results indicate that the footstep sounds generated by the proposed policy are quieter than those of the baseline. This finding suggests a correlation between the real-world footstep sound and the noisy walking penalties in the simulator. Consequently, these penalties such as foot contact velocity, joint acceleration, and base angular acceleration effectively enable aibo to walk quietly in the real world through our sim-to-real transfer approach.

C. Trade-off Between Robustness and Quietness

To assess robustness, we designed an experiment where aibo had to traverse a 0.5 m long slope within 20 seconds. We varied the slope angle and used this value as a metric to compare robustness, assuming that adaptability to unknown environments is one of the important aspect of robustness. We measured sound magnitude within the audible range using the

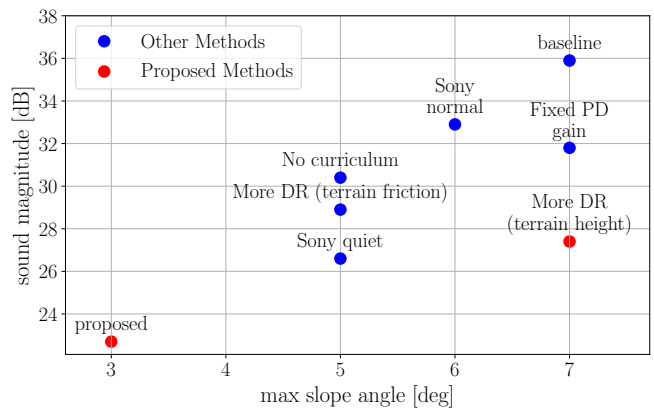


Fig. 4. Trade-off between robustness and quietness are shown. The experiment evaluates robustness by having aibo climb a slope. Sound magnitude is calculated using the same method as in Fig. 3. The controllers such as RL baseline, ablation test, SONY commercial controllers, and RL proposed are shown.

method described in subsection IV-A. We prepared policies including our proposed method, baseline, Sony commercial controllers, and ablation study conditions for the framework, which consists of three key factors.

Figure 4 shows that the loudest baseline policy can overcome the steepest slope of 7 degrees. Although our proposed method is the quietest among all settings, its robustness is the lowest. This observation suggests a trade-off between quietness and robustness, as quieter policies struggle to overcome steeper slopes.

We conducted an ablation study by training aibo in eight different settings. First, we evaluated policies without curriculum learning, using parameter scales from the noisy walking and quiet walking phases in Table II. As shown in Fig. 4 (No Curriculum), only using the noisy walking phase parameters without curriculum achieves about half of the noise reduction compared to the proposed method. Using the quiet walking phase parameters in Table II without curriculum, the training fails to converge. In other words, aibo doesn’t learn to walk, due to too severe penalties for quiet walking.

As a next ablation study, We tested the impact of removing switch contact sensors. Although the training converges, aibo doesn’t walk at all. At the beginning of the quiet walking phase, aibo falls frequently, ending episodes early. We hypothesize that this is due to the lack of information about when to stiffen joints to support the whole body. Consequently, aibo learns to avoid walking to satisfy the strong noisy walking penalties.

As a third ablation study, a comparable result in terms of robustness is obtained using our proposed method without allowing the network to change the PD gains, as shown in Fig. 4 (Fixed PD gain). While this demonstrates that the noisy walking penalties and curriculum can reduce noise, the sound magnitude differs from our proposed method. PD gain scale change is beneficial for quiet walking, as mentioned in previous work [29], because quiet walking is also a sensitive contact task.

Finally, we expanded our training of the proposed method

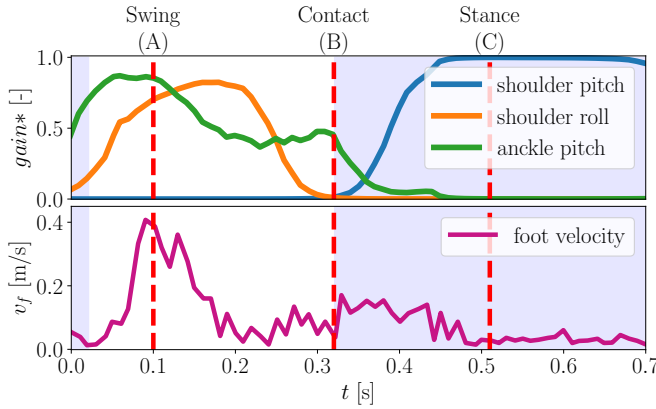


Fig. 5. PD gain scale of $\sigma(x)$ in the Eq. 1 at the fore right leg during locomotion. Right foreleg is (A): in the air to move forward, (B): approaching contact with the ground, (C): in contact to support the ensuing movement of the other legs. The plot at the bottom shows the edge of the foot velocity. The blue area in the plot shows when the right foreleg is in contact with the ground, based on the foot switch contact sensor.

by incorporating additional domain randomization parameters for ground characteristics, specifically the friction coefficient and terrain height. When modifying the friction coefficient range, increasing the maximum from 0.7 to 0.9 and decreasing the minimum from 0.4 to 0.2, we observed a relatively small sound magnitude but enhanced robustness compared to the originally proposed method as shown in Fig. 4 as More DR (terrain friction). Conversely, increasing the maximum terrain height from 0.01 to 0.03 resulted in the third-lowest sound magnitude while maintaining robustness comparable to the baseline as shown in Fig. 4 as More DR (terrain height). These findings suggest that the selection of domain randomization parameters enables tuning of the trade-off between robustness and quietness.

D. PD Gain Scale Analysis

Figure 5 illustrates the changes in PD gain scales and the edge of foot velocity exhibited by the quiet walking policy for the three actuators on the right foreleg. The plot of PD gain scale shows the scaled $\sigma(x_i)$ value outputted by the network, which is later scaled according to Equation 1.

As shown in the middle plot of Fig. 5, just before the foot makes contact with the ground, the gains for all the actuators of the right foreleg decrease, effectively dampening the motion. This drop in gain scales at the shoulder pitch joint is followed by an increase, supporting the ensuing movement of the other legs. This behavior suggests that the policy learns to walk quietly by adjusting the PD gains to dampen the joint at the moment of ground contact.

The foot velocity, calculated through forward kinematics using measurement data of the three joint angles for the right foreleg, is shown in Fig. 5. The foot velocity decreases just before contact with the ground. Conversely, when the foot is in the air, the velocity of the foot increases, along with the PD gains for shoulder roll and ankle pitch. This indicates that the policy actively controls the PD gains to maintain tracking of the base velocity.

V. DISCUSSION

While our proposed RL policy demonstrates improved quietness, its robustness is not as high as the baseline RL policy, resulting in reduced adaptability to unknown terrains. We have shown that the selection of DR parameters, such as ground friction and terrain height, can influence the trade-off between robustness and quietness of the policy. Future work could incorporate the approach of Chen et al. [39], which utilizes perception information to estimate ground characteristics like friction, height, and damping. This integration could enable the selection of an appropriate policy based on these factors, potentially allowing the robot to traverse slippery surfaces and employ robust walking strategies on various terrains.

Rather than directly minimizing footstep sound, which is challenging to emulate in physics simulator, we proposed sim-to-real RL framework to minimize foot contact velocity, which correlates with footstep sound. This approach could be extended to other applications where real-world parameters are difficult to emulate in the typical locomotion simulator. For instance, battery life optimization could be addressed by minimizing torque in the simulator, as torque correlates with energy consumption.

In this study, we focused primarily on reducing footstep sound during locomotion. However, other sources of noise during walking, such as actuator movement and mechanical friction, warrant further investigation. Although these sounds are not as dominant as footstep noise, addressing them remains crucial for developing quieter home-legged robots.

VI. CONCLUSION

In this work, we propose quiet walking based on a sim-to-real RL approach to minimize foot contact velocity, which is highly correlated with footstep sound for a small home robot. Our framework consists of three key elements: adaptive PD gains, utilization of foot contact sensors, and implementation of curriculum learning. Through ablation studies, we demonstrate that each element contributes to minimizing the footstep sound for achieving a quieter walk. From the perspective of sound magnitude in the audible range from 20 Hz to 20 kHz, our approach achieves the quietest locomotion controller compared to Sony’s commercial locomotion controllers and our own RL baseline. We also identify a trade-off between quietness and robustness.

Our work is the first to demonstrate and highlight key factors needed for the sim-to-real based RL to achieve quiet walking. These findings have important implications for future research in human-robot interaction and the commercialization of quadruped robots for home use.

ACKNOWLEDGMENT

The authors would like to thank Hiroyuki Izumi, Kensuke Kitamura, Ichitaro Kohara, Fumitaka Joo, Toshihisa Sambomatsu, Takuma Morita, and Yuichiro To at Sony Group Corporation for software integration help. Thanks to Yuntao Ma, Mayank Mittal and Jonas Frey at ETH Zurich for discussion about reinforcement learning.

REFERENCES

- [1] Sony, “aibo,” <https://us.aibo.com/>, online; accessed 30-Aug-2024.
- [2] Grove X, “Lovot,” <https://lovot.life/en/>, online; accessed 30-Aug-2024.
- [3] Yukai Engineering, “Qoobo,” <https://qoobo.info/index-en/>, online; accessed 30-Aug-2024.
- [4] G. Ruben, K. Espen, A. Michael, W. Georg, B. Jared, P. Steven, M. David, and B. Moritz, “Design and control of a bipedal robotic character,” *Robotics: Science and Systems (RSS)*, 2024.
- [5] B. Friedman, P. H. Kahn, and J. Hagman, “Hardware companions? what online aibo discussion forums reveal about the human-robotic relationship,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI ’03. New York, NY, USA: Association for Computing Machinery, 2003, p. 273–280.
- [6] K. Tanaka, H. Makino, K. Nakamura, A. Nakamura, M. Hayakawa, H. Uchida, M. Kasahara, H. Kato, and T. Igarashi, “The pilot study of group robot intervention on pediatric inpatients and their caregivers, using ‘new aibo,’” *European Journal of Pediatrics*, pp. 1–7, 2021.
- [7] G. F. Melson, P. H. Kahn, A. M. Beck, B. Friedman, T. Roberts, and E. Garrett, “Robots as dogs?: children’s interactions with the robotic dog aibo and a live australian shepherd,” in *CHI ’05 Extended Abstracts on Human Factors in Computing Systems*, 2005, p. 1649–1652.
- [8] H. S. Kang, K. Makimoto, R. Konno, and I. S. Koh, “Review of outcome measures in paro robot intervention studies for dementia care,” *Geriatric Nursing*, vol. 41, no. 3, pp. 207–214, 2020.
- [9] V. Makovychuk, L. Wawrzyniak, Y. Guo, M. Lu, K. Storey, M. Macklin, D. Hoeller, N. Rudin, A. Allshire, A. Handa, and G. State, “Isaac gym: High performance GPU based physics simulation for robot learning,” in *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*, 2021.
- [10] J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter, “Learning quadrupedal locomotion over challenging terrain,” *Science Robotics*, vol. 5, no. 47, p. eabc5986, 2020.
- [11] T. Miki, J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter, “Learning robust perceptive locomotion for quadrupedal robots in the wild,” *Science Robotics*, vol. 7, no. 62, p. eabk2822, 2022.
- [12] S. Choi, G. Ji, J. Park, H. Kim, J. Mun, J. H. Lee, and J. Hwangbo, “Learning quadrupedal locomotion on deformable terrain,” *Sci Robot*, vol. 8, no. 74, p. eade2256, Jan. 2023.
- [13] J. Wu, G. Xin, C. Qi, and Y. Xue, “Learning robust and agile legged locomotion using adversarial motion priors,” *IEEE Robotics and Automation Letters*, vol. 8, no. 8, pp. 4975–4982, Aug. 2023.
- [14] F. Jenelten, J. He, F. Farshidian, and M. Hutter, “DTC: Deep tracking control,” *Science Robotics*, vol. 9, no. 86, p. eadh5401, 2024.
- [15] N. Rudin, D. Hoeller, P. Reist, and M. Hutter, “Learning to walk in minutes using massively parallel deep reinforcement learning,” in *Proceedings of the 5th Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, A. Faust, D. Hsu, and G. Neumann, Eds., vol. 164. PMLR, 08–11 Nov 2022, pp. 91–100.
- [16] Y. Yang, T. Zhang, E. Coumans, J. Tan, and B. Boots, “Fast and efficient locomotion via learned gait transitions,” in *Proceedings of the 5th Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, A. Faust, D. Hsu, and G. Neumann, Eds., vol. 164. PMLR, 08–11 Nov 2022, pp. 773–783.
- [17] A. Escontrela, X. B. Peng, W. Yu, T. Zhang, A. Iscen, K. Goldberg, and P. Abbeel, “Adversarial motion priors make good substitutes for complex reward functions,” in *IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*, 2022, pp. 25–32.
- [18] Y. Fuchioka, Z. Xie, and M. Van de Panne, “Opt-mimic: Imitation of optimized trajectories for dynamic quadruped behaviors,” in *IEEE Int. Conf. Robot. Autom. (ICRA)*, 2023, pp. 5092–5098.
- [19] A. Reske, J. Carius, Y. Ma, F. Farshidian, and M. Hutter, “Imitation learning from mpc for quadrupedal multi-gait control,” in *IEEE Int. Conf. Robot. Autom. (ICRA)*, 2021, pp. 5014–5020.
- [20] E. Parizet, E. Guyader, and V. Nosulenko, “Analysis of car door closing sound quality,” *Applied acoustics*, vol. 69, no. 1, pp. 12–22, 2008.
- [21] H. Tennent, D. Moore, M. Jung, and W. Ju, “Good vibrations: How consequential sounds affect perception of robotic arms,” in *2017 26th IEEE international symposium on robot and human interactive communication (RO-MAN)*, 2017, pp. 928–935.
- [22] G. Trovato, R. Paredes, J. Balvin, F. Cuellar, N. B. Thomsen, S. Bech, and Z.-H. Tan, “The sound or silence: investigating the influence of robot noise on proxemics,” in *2018 27th IEEE international symposium on robot and human interactive communication (RO-MAN)*, 2018, pp. 713–718.
- [23] F. Asano, “Stealth walking of 3-link planar underactuated biped,” in *IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*, 2017, pp. 4118–4124.
- [24] M. W. Gomes and K. Ahlin, “Quiet (nearly collisionless) robotic walking,” *IEEE Int. Conf. Robot. Autom. (ICRA)*, pp. 5761–5766, 2015.
- [25] D. Moore, N. Martelaro, W. Ju, and H. Tennent, “Making noise intentional: A study of servo sound perception,” in *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2017, pp. 12–21.
- [26] B. J. Zhang, K. Peterson, C. A. Sanchez, and N. T. Fitter, “Exploring consequential robot sound: Should we make robots quiet and kawaii-et?” in *IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*, 2021, pp. 3056–3062.
- [27] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” *arXiv preprint arXiv:1707.06347*, 2017.
- [28] X. B. Peng and M. Van De Panne, “Learning locomotion skills using deeprl: Does the choice of action space matter?” in *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, 2017, pp. 1–13.
- [29] M. Bogdanovic, M. Khadiv, and L. Righetti, “Learning variable impedance control for contact sensitive tasks,” *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6129–6136, 2020.
- [30] K. Ohnishi, M. Shibata, and T. Murakami, “Motion control for advanced mechatronics,” *IEEE/ASME transactions on mechatronics*, vol. 1, no. 1, pp. 56–67, 1996.
- [31] M. Umair Bin Altaf, T. Butko, and B.-H. Juang, “Acoustic gaits: Gait analysis with footstep sounds,” *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 8, pp. 2001–2011, 2015.
- [32] W. Yu, G. Turk, and C. K. Liu, “Learning symmetric and low-energy locomotion,” *ACM Trans. Graph.*, vol. 37, no. 4, jul 2018.
- [33] J. Tuyls, S. Yao, S. Kakade, and K. Narasimhan, “Multi-stage episodic control for strategic exploration in text games,” *arXiv preprint arXiv:2201.01251*, 2022.
- [34] A. Hartmann, D. Kang, F. Zargarbashi, M. Zamora, and S. Coros, “Deep compliant control for legged robots,” in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 11 421–11 427.
- [35] J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel, “Domain randomization for transferring deep neural networks from simulation to the real world,” in *IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*, 2017, pp. 23–30.
- [36] Y. Song, A. Romero, M. Müller, V. Koltun, and D. Scaramuzza, “Reaching the limit in autonomous racing: Optimal control versus reinforcement learning,” *Science Robotics*, vol. 8, no. 82, p. eadg1462, 2023.
- [37] P. Dale and M. W. Stephen, “Neuroscience. 2nd edition,” no. 2, 2001.
- [38] P. Welch, “The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms,” *IEEE Transactions on audio and electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.
- [39] J. Chen, J. Frey, R. Zhou, T. Miki, G. Martius, and M. Hutter, “Identifying terrain physical parameters from vision-towards physical-parameter-aware locomotion and navigation,” *arXiv preprint arXiv:2408.16567*, 2024.