

Towards Standardized Disturbance Rejection Testing of Legged Robot Locomotion with Linear Impactor: A Preliminary Study, Observations, and Implications

Bowen Weng¹, Guillermo A. Castillo², Yun-Seok Kang³, and Ayonga Hereid⁴

Abstract—Dynamic locomotion in legged robots is close to industrial collaboration, but a lack of standardized testing obstructs commercialization. The issues are not merely political, theoretical, or algorithmic but also physical, indicating limited studies and comprehension regarding standard testing infrastructure and equipment. For decades, the approaches we have been testing legged robots were rarely standardizable with hand-pushing, foot-kicking, rope-dragging, stick-poking, and ball-swinging. This paper aims to bridge the gap by proposing the use of the linear impactor, a well-established tool in other standardized testing disciplines, to serve as an *adaptive, repeatable, and fair* disturbance rejection testing equipment for legged robots. A pneumatic linear impactor is also adopted for the case study involving the humanoid robot Digit. Three locomotion controllers are examined, including a commercial one, using a walking-in-place task against frontal impacts. The statistically best controller was able to withstand the impact momentum ($26.376 \text{ kg} \cdot \text{m/s}$) on par with a reported average effective momentum from straight punches by Olympic boxers ($26.506 \text{ kg} \cdot \text{m/s}$). Moreover, the case study highlights other anti-intuitive observations, demonstrations, and implications that, to the best of the authors’ knowledge, are first-of-its-kind revealed in real-world testing of legged robots.

I. INTRODUCTION

Dynamic locomotion, i.e. the capability to create statistically stable and adaptive walking gaits that can withstand a certain degree of external disruption, has been a central pursuit for legged robot researchers for many decades [1], [2], [3], [4], [5], [6]. The complex, articulated structure of mechanical limbs, coupled with the prevalent under-actuation and the uncertainties encountered in real-world environments, present a distinct and critical challenge in the field of legged robotics. This challenge distinguishes legged robots from other mobility alternatives, such as wheeled robots, in its uniqueness and complexity.

Fortunately, the past decades have also observed a tremendous amount of proposals tackling the locomotion challenge against external disturbances [2], [3], [4], [5], [7], [8], [6], [9]. Many of the proposals also demonstrate real-world

operations or in the wild. However, there still is a large gap towards scaling up the deployment of legged robots to work alongside humans in an array of industrial and commercial environments. One of the main bottlenecks lies with the lack of formal and unbiased performance testing standards and certificates. To a very large extent, even under specific functionalities such as the disturbance rejection domain with fallover being the only failure event of concern, one does not necessarily know if the robot “works”, “works well”, or “works well enough”. The remediate of this gap requires a series of cross-disciplinary work addressing the theoretical foundation of testing [10], the testing algorithms [11], and the development of testing equipment and infrastructure [12]. In particular, the last topic has been significantly overlooked and understudied in the legged robot field and is also the primary focus of this paper.

For the past few decades, *researchers have been largely relying on non-rigorous, non-repeatable, and occasionally unfair testing apparatus for the conceptual performance demonstration of legged robots against external disturbances*. We used hand-pushing [6], foot-kicking, rope-dragging [9], hockey sticks poking [13], yoga-ball-throwing [9], and swing-ball impact [5]. Despite their widespread use, intuitive (and somewhat impressive) demonstrations, these approaches are barely a solid test of the robots’ true capabilities as they lack standardization, consistency, and rigorous flexibility. The researchers have also performed simulation-oriented tests generating the external impact as forces posed upon a certain contact point on the robot [6], [11]. Those forces are often applied to the robot in an unrealistic manner. For example, the direction of the external force often remains still regardless of the orientation and position variations of the robot body during the contacting period after impact. The forces are also configured to exhibit some “ideal” patterns, such as Quasi-static, periodic, and spiking patterns. All of the above require accurate control of the impact time and impact duration, which would not be possible in real-world practice against intelligent legged robots, and could also lead to unfair tests as one will clarify later in Section IV.

A. Main Contributions

This paper introduces a *novel proposal to utilize an established piece of impact generation equipment*, specifically referred to as the *linear impactor* [14], [15], [16], [17], for the standardized testing of legged robot locomotion

This work was supported in part by the National Science Foundation under grant FRR-21441568 and in part by National Highway Traffic Safety Administration (NHTSA) U.S. Department of Transportation.

¹Department of Computer Science, Iowa State University, Ames, IA, USA; bweng@iastate.edu.

²Electrical and Computer Engineering, The Ohio State University, Columbus, OH, USA; castillomartinez.2@osu.edu.

³School of Health and Rehabilitation Sciences, The Ohio State University, Columbus, OH, USA. yunseok.kang@osumc.edu.

⁴Mechanical and Aerospace Engineering, The Ohio State University, Columbus, OH, USA. hereid.1@osu.edu.

against external physical disturbances such as impacts¹. This adaptation positions the linear impactor as a candidate for creating rigorous, objective, and repeatable performance tests, enabling the assessment of legged robots in a manner that is both fair and adaptable. To our knowledge, this study represents the first effort within the field of legged robotics to address the need for suitable testing equipment and infrastructure, focusing on the aspects that ensure unbiased and consistent evaluation.

Moreover, a *real-world in-lab experiment* was conducted to assess the disturbance rejection performance of the Digit humanoid robot. This involved using a pneumatic linear impactor to subject the robot, equipped with three different state-of-the-art locomotion controllers, to various physical disturbances. Notably, this same pneumatic impactor was previously employed in studies on injury biomechanics [16], head and helmet impacts [15], underscoring its utility across diverse applications. The experiment unveiled several unique insights and challenges specific to the testing of legged robots. Among these findings was the anti-intuitive observation that a more severe impact does not necessarily translate into a worse outcome. This is attributable to the complex and intelligent feedback mechanisms inherent in legged robot locomotion control, a phenomenon that had only been identified in simulation-based studies until now [11], [10].

II. PRELIMINARIES AND PROBLEM FORMULATION

A. Problem Formulation

The testing subject in this study is a certain legged robot capable of achieving the functionality of dynamic locomotion. It essentially admits the system formulations as

$$\mathbf{s}(t+1) = h(\mathbf{s}(t), \mathbf{u}(t); \omega(t)). \quad (1)$$

with state $\mathbf{s} \in S \subset \mathbb{R}^n$, action $\mathbf{u} \in \mathcal{U} \subset \mathbb{R}^m$, and uncertainties $\omega \in W \subset \mathbb{R}^w$. The discrete-time characterization suits the practical needs with testing execution and data acquisition primarily involving digital equipment. Note the uncertainties are not involving any environmental factors, but are primarily internal such as the varying frictions of joints, the latency in the controller, and sensor noise. Note this approach differs from a typical stochastic system formulation for legged robots. In the context of testing, the so-called environmental disturbances are not randomly presented but are explicitly given and controlled with precision. This deliberate manipulation contrasts with traditional methods, allowing for more accurate assessments and understandings of the system's response to specific disturbances.

As a result, the control action \mathbf{u} includes two major components as $\mathbf{u} = [\mathbf{u}_i \ \mathbf{u}_e]$. The \mathbf{u}_i denotes the typical control signal that is part of the testing subject such as the commended torques at all actuators (hence the subscript “i” highlights the “internal” nature). On the other hand, \mathbf{u}_e denotes the environmental factors and external disturbances such as the wind speed, the push-over action, the presence of other obstacles, and the room temperature. Those are

the essential factors to be controlled in the context of performance testing.

As the focus of this paper is the disturbance rejection performance against external impact, the primary component of \mathbf{u}_e is thus a certain notion of the physical “impact”. Other factors will also be rigorously controlled and monitored to ensure they do not pose significant effects on the accuracy and repeatability of the testing outcomes.

In the practice of tests, the impact action \mathbf{u}_e is typically generated through a certain testing equipment operated by human engineers including the physical apparatus and other control and monitoring software modules. Formally speaking, we have the determination of \mathbf{u}_e satisfying

$$\mathbf{u}_e(t) = g(\mathbf{x}(t), \mathbf{a}(t); \omega_g(t)), \quad (2)$$

with state $\mathbf{x} \in X \subset \mathbb{R}^x$, action $\mathbf{a} \in \mathcal{A} \subset \mathbb{R}^a$, and uncertainties $\omega_g \in W_g \subset \mathbb{R}^{\omega_g}$. The \mathbf{a} denotes the direct action one can take as a human testing operator (e.g., the releasing height of the pendulum impactor and the air pressure of the pneumatic impactor) for controlling the testing equipment. The state space $X \subset S$ as it is primarily concerned with the subject features directly related to the testing action propagation (e.g., the specific status of joints and actuators are not of primary concern for the impact testing equipment hence they are not part of \mathbf{x}).

Moreover, the internal control actions are determined by a certain locomotion control policy

$$\mathbf{u}_i(t) = \pi(\mathbf{s}(t)). \quad (3)$$

Composing (3) with (1), one has

$$\begin{aligned} \mathbf{s}(t+1) &= h(\mathbf{s}(t), [\pi(\mathbf{s}(t)) \ \mathbf{u}_e(t)]; \omega(t)) \\ &= f(\mathbf{s}(t), \mathbf{u}_e(t); \omega(t)). \end{aligned} \quad (4)$$

Note the testing also involves the black-box system nature [10] with $f \in \mathcal{F}$ and $\mathcal{F} = S^{S \times \mathcal{U}_e \times W}$ being the set of all possible mappings. This has also been concretely observed in practice given the variety of legged robot hardware mechanism designs and software controller development. Note the unknowability and the stochasticity are both part of the subject nature and should be respected by the testing perspective.

The purpose of this paper is to present a certain testing equipment with system g as defined by (2) such that it is

- 1) **Adaptive:** The generated \mathbf{u}_e works with a variety of systems and robots $f \in \mathcal{F}$;
- 2) **Repeatable:** The system g satisfies $W_g = \emptyset$ and for all $\mathbf{x} \in X$ and any pair of sufficiently close testing actions, \mathbf{u}_e^1 and \mathbf{u}_e^2 with $\|\mathbf{u}_e^1 - \mathbf{u}_e^2\|_p$ sufficiently small, $\|g(\mathbf{x}, \mathbf{u}_e^1) - g(\mathbf{x}, \mathbf{u}_e^2)\|_p$ is also sufficiently small;
- 3) **Fair:** \mathbf{u}_e is strictly determined by g for fixed state-actions, and should not be impacted by h or f .

Unfortunately, existing impact generating solutions with hands, feet, hockey sticks, and yoga balls for legged robot testing rarely satisfy the above properties. This paper is thus inspired to explore impact generation in a broader scope with lessons learned from other disciplines.

¹Related video: <https://youtu.be/bUWnnTYNL3A>

B. Creating the Impact

The notion of creating a physical “impact” and identifying outcomes thereafter has been a commonly observed testing configuration throughout multiple disciplines with materials [18], biomechanics [16], [15], consumer electronics [19], helmets [20], to name a few.

One of the earliest impact tests is the Charpy Impact Test, also known as the Charpy V-notch test, Notched Bar Impact Testing, and Pendulum Impact Test, was first proposed in 1896 [18], to determine the amount of energy absorbed by a material during fracture. It has further extended to a variety of tests against helmets, footwear, electronics, etc. The impact was primarily generated with a pendulum of known mass and length that is dropped from a predetermined height. A similar setup called the Ball Impact Test (i.e. replacing the pendulum with a ball attached to a swigging string) has also been widely adopted in the glass industry to determine the strength and integrity of glass.

A significant limitation of the experimental setup described above, particularly when applied to the testing of legged robots, lies in its lack of flexible adaptability. Traditional testing methods, such as those employing pendulum impacts or swing-ball mechanisms, are often constrained by their reliance on gravity and the specific physical setup. These factors limit the impact momentum and speed that can be achieved, and are mostly comparable with small-scale and stationary subjects. Such gravity-dependent testing mechanisms present substantial challenges when attempting to adapt them to more complex scenarios, such as interactions with large-scale or moving subjects. The inherent design limitations restrict the ability to achieve impact motions with higher degrees of freedom, thereby reducing the test’s relevance to real-world scenarios where multi-dimensional movements and forces are encountered. The same problem has also been observed with other impact generators such as the Shock Testing and Drop Tower Tests [21], [22].

III. TESTING LEGGED ROBOT LOCOMOTION WITH THE LINEAR IMPACTOR

In this section, the linear impactor is introduced as a proposed testing equipment towards a standardized disturbance rejection testing of legged robots.

A. The Linear Impactor

Intuitively, a linear impactor produces the impact motion by driving a certain object (commonly referred to as the impactor) along a linear path to impact the subject target. One of the common ways to initiate the impact is through compressed air with the corresponding impactor referred to as the pneumatic impactor. It has gained extensive applications in a variety of disciplines. The pneumatic impactor used in this study (see Fig. 1) has also been traditionally adopted as the standard testing equipment to support research activities related to vehicle occupant safety, injury biomechanics, and sports biomechanics [16], [15]. In practice, there are other designs such as the Hydraulic Impactor [23], Mechanical

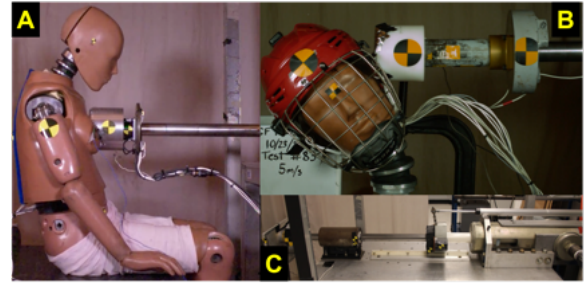


Fig. 1: The linear pneumatic impactor has been used as a standard testing equipment related to the research in automobile safety and injury biomechanics simulating impact energy to head, tibia, thorax, abdomen, and shoulder impacts: A shows the test of dummy thorax impact, B is coming from the study of head and hockey helmet instrumentation evaluation [15], both tests were performed with the linear pneumatic impactor with different impact energies shown in C. The same device is also configured with minor modifications for legged robot locomotion testing performed in this study.

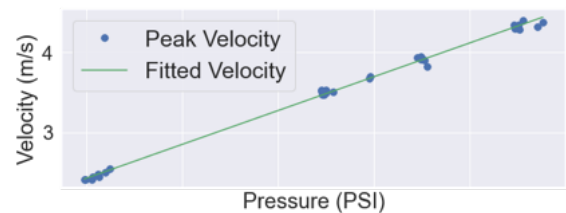


Fig. 2: The pneumatic impactor adopted in this study demonstrates its repeatability by empirically showing a linear relation between the testing action u_e , the peak velocity achieved by the impactor, and the operator control a , the configured pressure for the compressed air. The linear fitting error is within 0.1 m/s.

Impactor, Motor-Driven Impactor, and Electromagnetic Impactor [24], some of which are not necessarily confined to the testing discipline [23].

The setup of the linear impactor allows the testing subject to be configured in an open space with sufficient room for extended motion. The impactor face is free to deflect and rebound in a natural way to assist the study of the complete impact kinematics. A variety of ram shafts, components, and impactor faces can be configured to achieve desired impact energies and simulated impact surfaces. The impactor can also be installed on other mobile platforms to achieve a higher degrees-of-freedom. Its *adaptivity* is hence immediate given the flexible configurations mentioned above.

B. Testing Configuration and Repeatability

The primary class of subject robot being tested in this paper is the class of two-legged robots including the bipedal and the humanoids. Yet the methodologies can also be extended to other types of legged robots such as the quadruped.

The robot could be configured with different designs revealed by the configuration of (1) and equipped with different locomotion algorithms formulated as (3), leading to a variety of testing subject. They are then tested with the testing control being the impact as described by (4). Every “test” thus involves a series of manipulations centered around the presented dynamics of (4).

Generally speaking, every test starts by configuring the robot at a certain state $s(0) = s_0 \in S$. For the walking-in-place example, s_0 characterizes the position, the desired

walking gaits with zero velocity, and the orientation of the robot. As highlighted with (1), the specific stance phase, swing phase, swing foot clearance at the moment of impact are essentially part of the uncertainties ω . The diversities of those uncertainties should be respected by the testing program and it would be unfair to force the robot to be tested at a specific walking phase.

Reflected in the practice of testing, the robot needs to be brought or controlled into the desired s_0 or a sufficiently small neighbourhood near s_0 . This is often with the assistance of a human testing operator. Note the human or external assistance is a common setup in the practice of testing. For example, the testing of Automated Driving Assist Systems also requires the expert testing driver or a so-called steering robot to bring the vehicle to a certain desired speed before starting the testing procedure [25]. The impact event is further initiated by configuring the operator action a (e.g. the compressed air pressure for the pneumatic impactor used in the next section). The impact action u_e is then propagated through the impactor equipment.

Note the impact action in this study is the maximum velocity the impactor manages to achieve before the impact referred to as the *peak velocity*. For the testing subject that is stationary, one can align the impact velocity (the velocity of the impactor at the moment of impact) with the peak velocity sufficiently well by placing the subject appropriately. This is not possible for the legged robot testing. The torso of the robot cannot remain absolutely still and the robot cannot stay “absolute” in-place even the desired velocity is set to zero. The variances exhibited in the walking patterns could be of significant diversities among the testing subject. They are components of the internal uncertainties characterized by ω in (1) and could also be the choice of the internal control signal u_i determined by each individual locomotion controller (3). The position drifts and torso shaking as mentioned above are empirically minor throughout all tests performed in this study (see Fig. 6 and Fig. 7 for example), yet their impact to the locomotion performance outcome is unclear.

This section is concluded with the repeatability demonstration (see Fig. 2) of the linear pneumatic impactor mentioned above and used throughout all tests discussed in the next section. The repeatability is demonstrated within the velocity range between 1.5 m/s and 4.4 m/s with the error of less than 0.1 m/s. The particular impactor is capable of reaching approximately 10 m/s velocity with the similar repeatability performance, yet that are beyond the current capability of the subject robot studied in this paper.

IV. A CASE STUDY WITH PNEUMATIC IMPACTOR AND WALKING-IN-PLACE TESTING

For the particular pneumatic impactor adopted in the case study section discussed next, the overall configuration is shown in Fig. 3. The center of the impactor face is 99.2 cm from the ground. The impactor weights 6.4 kg. The impactor face is a rectangle of 6-inch \times 4-inch (about 15.24 cm \times 10.16 cm).

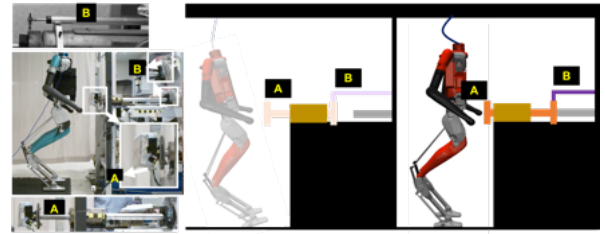


Fig. 3: The testing configuration in-lab and the conceptual illustration before and after the impact: “A” denotes the impactor face with a rectangular surface, “B” denotes the displacement potentiometer, one can also refer to Fig. 1C for another view of the same equipment.

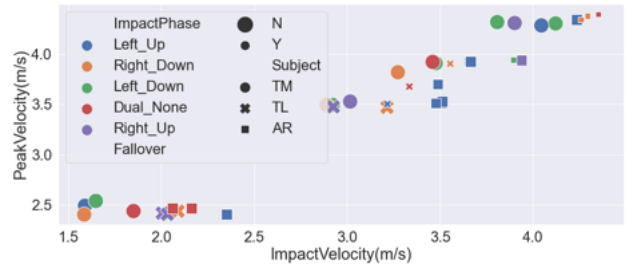


Fig. 4: An overview of all 36 impact tests against three locomotion controllers: different impact phases are differentiated by the marker color and the label is specified in the form of stance foot phase (whether the left or the right foot is on the ground) and swing foot phase (whether the swing foot is going up or stepping down) with an underscore “_” in between. Note the dual support phase (with both feet on-ground) is specified as “Dual_None”. The fallover status is characterized by the size of the marker. The three different locomotion controllers are differentiated by the marker type.

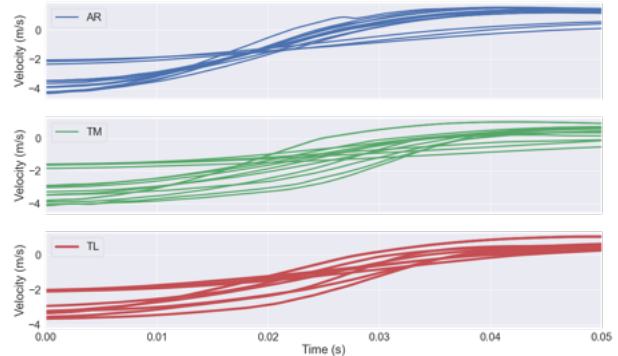


Fig. 5: The velocity profiles of the impactor within 0.05-second after the moment of impact.

A. Subject Robot Configuration

The same Digit robot is adopted throughout all impact tests. Three different locomotion controllers are tested, including (i) a template-model (TM) inspired whole body controller, (ii) a template-model based task space learning policy (TL) [26], and (iii) a previous version of the default controlled coming with Digit designed by its original manufacturer Agility Robotics (AR). Specifically, the TM controller is primarily based on the Task Space Inverse Dynamics (TSID) methodology introduced in [27] with a high-level step planner based on Raibert’s regulations [28]. The TL controller follows the hierarchical framework in [26] that combines a reinforcement learning-based high-level planner policy for the online generation of task space commands with a low-level TSID controller to track the desired task space trajectories. The AR controller is part of a commer-

cially available product, hence the technical details are not open and unclear. Those controllers are also different at the selected gait (e.g., stepping frequency and swing foot clearance), the desired torso height, and many other known, unknown, observable, and non-observable configurations. All of the controllers have demonstrated an overall good performance for stepping-in-place, disturbance rejection, walking over various terrain conditions, and in the wild. However, it remains unclear to what concrete extent the locomotion control is stable and safe (i.e. performance characterization), which one is the best (i.e. benchmark comparison), to name a few.

B. Testing Configuration

For start, as discussed in Section III, the robot is controlled to achieve the desired walking-in-place motion at a fixed point near the impactor where the head is expected to reach its peak velocity. For TM and TL, the robot is brought to the fixed point with the assistance of the human testing operator and the walking-in-place is triggered by configuring the desired velocity to zero (yet one may still drift as mentioned above in Section III). For AR, the controller is technically not capable of walking-in-place as the robot will simply stop and stand still if the the desired velocity is set to zero. As a result, the walking-in-place motion for AR is achieved with the human operator's help through consecutively giving small desired speed signals (≤ 0.1 m/s) to keep the robot close to the fixed point. For all impact tests performed, the robot remains within a short distance near the fixed point. This can be demonstrated by the short time gaps between the peak velocity and the identified impact velocity as shown in Fig. 6 and Fig. 7. Specifically, among all 36 tests, the average time gap between reaching the peak velocity and the moment of impact is $0.0453 (\pm 0.0342)$ s.

To help protect the robot, especially the sensors (one of the camera modules is near the estimated impact point), a Polyethylene foam block of 5.4 cm thickness are installed at the expected impact point and an Ethafoam block is installed on the impactor face. This can be viewed as part of the system (1) and remains configured in the same way for all tests against all subject controllers. As shown in Fig. 3, the robot is also loosely connected to a sliding rail installed above the robot. This is an added protection to the robot to prevent mechanical breaks and other damages caused by an actual fallover. As a result, the robot will not physically fallover (i.e. having any part of the robot physically touching the ground other than the two feet). The justification of whether a robot fails in the test is justified by combining the testing operator's observation and the post analyses of the dynamic states of robots after each test.

For each subject locomotion controller, the tests start with the controlled pressure around 50 PSI (Pounds per square inch) and one gradually increases the pressure to up to 95 PSI. If the subject controller fails (i.e., falls over) consecutively two trials, the tests will stop and one moves on to the next subject to prevent potential damages to the robot from repetitive failures. Note real-world testing is a complex

process. The initialization, calibration, reset, configuration, and many other standard procedures highlight the challenges of performing rigorous real-world tests which may have been overlooked by some existing works in the field. The 36 impact events are conducted for the purpose of a preliminary study with the proposal of this paper. Despite the limited number of tests, we can already draw some interesting observations and insights, which will be discussed next.

C. Observations and Analyses

An overview of the 36 impact events are summarized in Fig. 4. The stepping phase at the moment of impact, the fallover outcome, and the subject controller involved in the test are also illustrated. There are two specific observations/insights one shall highlight in the section: (i) the necessity of characterizing the impact testing action \mathbf{u}_e as impact momentum (instead of force) to ensure the testing fairness, and (ii) the anti-intuitive observation where a more significant impact does not necessarily lead to a more severe outcome.

First, in the legged robot regime, the external force has been adopted extensively to characterize the impact disturbance, especially in the computer simulations [29], [30], [6], [26], [11]. This may not be a fair setup as the impact force is jointly determined by the impact momentum and impact duration. The impact duration is a changing variable that is not only determined by the testing equipment/algorithm, but more importantly, by the subject itself. As shown in Fig. 5 (also partially revealed in Fig. 6 and Fig. 7), throughout all tests, the impactor's velocity exhibits a similar pattern after the moment of impact, i.e., decaying to zero and sometimes bouncing back (with a positive velocity). However, the impactor velocity patterns in between different subject controllers are also quite different. For example, comparing TM with AR in Fig. 5, the impactor decelerates faster in the AR cases than that in the TM ones, especially for large impact velocities. A primary cause for these differences is the design of the locomotion controller, specifically how it orchestrates its periodic motion, manages feedback control for walking-in-place, and responds to disturbances. From the perspective of the testing operator, the intricacies of these designs remain largely undisclosed. Consequently, they might display significantly varied behaviors, resulting in different impact durations. These durations, however, cannot be directly used as clear indicators of performance. As a result, it is only fair to take the impact momentum, i.e. the momentum the testing equipment generates at the moment of impact, as the testing action \mathbf{u}_e . It may also be different from the repeatably controlled peaking velocity as the legged robot cannot maintain absolute in-place as mentioned above, which is also part of the robot's nature in design. Among the three locomotion controllers tested, the most effective one (TM) was able to withstand an impact momentum of $26.376 \text{ kg} \cdot \text{m/s}$. Intriguingly, this figure is roughly on par with the average effective momentum of straight punches delivered by Olympic boxers, which is reported to be $26.506 \text{ kg} \cdot \text{m/s}$ [31]. Although the details are not rigorously controlled and com-

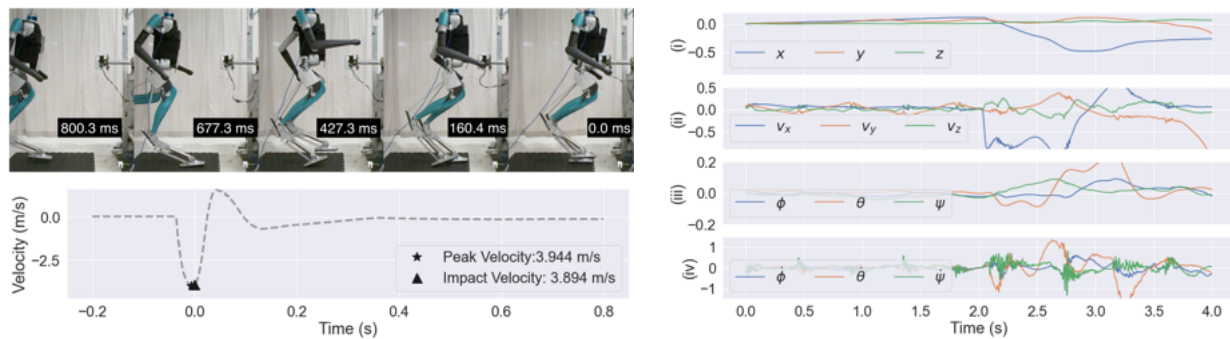


Fig. 6: The AR-equipped Digit robot falls over against the frontal impact velocity at 3.89 m/s: the upper left sub-figure denotes the frames captured by the high-speed camera at and after the moment of impact (from right to left), the lower left sub-figure illustrates the velocity profile of the impactor within the time range in between 0.2 s before the impact and 0.8 s after the impact, the right sub-figure shows the robot dynamic states in a 4-second window including the impact moment, the first row denotes center of mass (CoM) position of the torso, the second row denotes the CoM velocity, the third row shows the roll (ϕ), pitch (θ), and yaw (ψ) angles of the robot torso, and the last row shows the roll rate, pitch rate, and yaw rate of the robot torso.

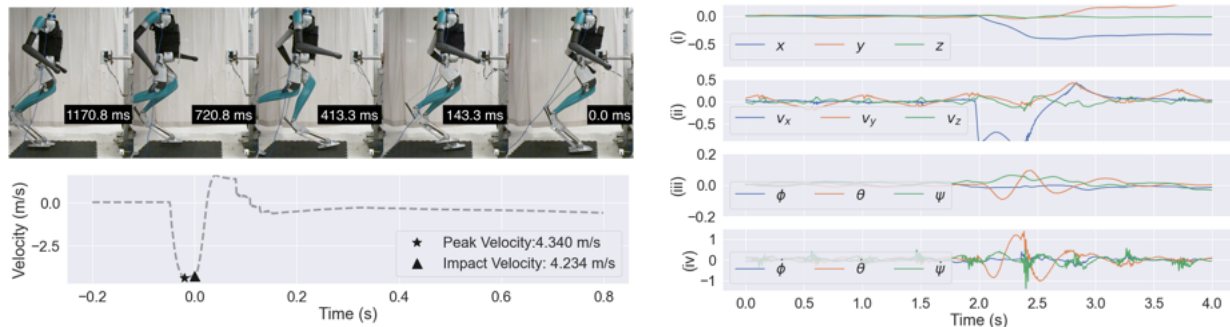


Fig. 7: The AR-equipped Digit robot recovers from the frontal impact velocity at 4.23 m/s (a more significant impact compared with Fig. 6): the sub-figures share a similar setup as described in the caption of Fig. 6.

parable such as the impact face materials, this conceptual comparison may foster new perspectives on evaluating the performance of robotic systems in terms of “how safe is safe enough”. Specifically, it introduces the possibility of drawing measurable equivalences between the performance of legged robots and human biomechanics, contributing to a more nuanced understanding of what constitutes an adequate level of safety in legged robotic applications against external disturbances. For the second observation mentioned above, as one can see from Fig. 4, there are multiple regions where impact events sharing similar impact velocities end up with different fallovers (indicated by the marker size). For example, the TL-driven robot recovers from the impact velocities of 3.214 m/s and falls over against the 3.216 m/s velocity, and that is only 0.0128 kg·m/s difference in impact momentum. A similar pattern is seen with AR at notably higher impact velocities: 3.89 m/s, 4.23 m/s, and 4.26 m/s. Out of these three impacts, the robot falls over twice, but intriguingly, the two falls aren’t associated with the highest velocities. Detailed visuals related to two of the three AR impacts are available in Fig.6 and Fig.7. It’s worth noting that while the reason behind this phenomenon is not our primary concern, the existence of such counter-intuitive cases emphasizes the challenges inherent in creating equitable and effective safety performance testing algorithms [10].

V. SUMMARY AND FUTURE WORK

This paper has presented a proposal of adapting the well-adopted impact generation equipment, namely the *linear impactor*, as a candidate testing equipment to achieve the

standardized performance testing of legged robot locomotion against external physical impact. We’ve also conducted initial laboratory tests with a linear pneumatic impactor on the Digit humanoid robot with three different locomotion controllers. These experiments have revealed interesting insights, challenges, and lessons, all contributing to an expected future of standardized safety performance testing for legged robots. Given this being a preliminary study, there are multiple directions of future interest including rigorous statistical analyses with increased tests, the expansion to other locomotion tasks [12], the cyber-physical integration with advanced testing algorithms [11], [32], and other enhancement of the linear impactor that better adapts to the specific needs of legged robot testing under various conditions.

ACKNOWLEDGMENT

The authors thank the assistance from the students and staff at Injury Biomechanics Research Center and Cyberbotics Lab members at The Ohio State University (OSU), including Angelo Marcallini, Rosalie Connell, Zachary Haverfield, Sankalp Agrawal, etc., and the support from Michelle Murach at Transportation Research Center Inc. (TRC) and Dr. Kevin Moorhouse from National Highway Traffic Safety Administration (NHTSA). Positions and opinions advanced in this work are those of the authors and not necessarily those of OSU, TRC, or NHTSA. Responsibility for the content of the work lies solely with the authors. The authors also thank the insightful comments and suggestions from anonymous ICRA reviewers and editors (though the 6-page limit is too restrictive to have everything addressed).

REFERENCES

- [1] D. G. Hobbelen and M. Wisse, "A disturbance rejection measure for limit cycle walkers: The gait sensitivity norm," *IEEE Transactions on Robotics*, vol. 23, no. 6, pp. 1213–1224, 2007.
- [2] H. Diedam, D. Dimitrov, P.-B. Wieber, K. Mombaur, and M. Diehl, "Online walking gait generation with adaptive foot positioning through linear model predictive control," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2008, pp. 1121–1126.
- [3] M. Posa, T. Koolen, and R. Tedrake, "Balancing and step recovery capturability via sums-of-squares optimization," in *Proceedings of Robotics: Science and Systems*, Cambridge, Massachusetts, July 2017.
- [4] A. Meduri, P. Shah, J. Viereck, M. Khadiw, I. Havoutis, and L. Righetti, "Biconmp: A nonlinear model predictive control framework for whole body motion planning," *IEEE Transactions on Robotics*, vol. 39, no. 2, pp. 905–922, 2023.
- [5] H. Chen, Z. Hong, S. Yang, P. M. Wensing, and W. Zhang, "Quadruped capturability and push recovery via a switched-systems characterization of dynamic balance," *IEEE Transactions on Robotics*, vol. 39, pp. 2111–2130, 2023.
- [6] G. A. Castillo, B. Weng, W. Zhang, and A. Hereid, "Robust feedback motion policy design using reinforcement learning on a 3d digit bipedal robot," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 5136–5143.
- [7] 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.
- [8] J. Hwangbo, J. Lee, A. Dosovitskiy, D. Bellicoso, V. Tsounis, V. Koltun, and M. Hutter, "Learning agile and dynamic motor skills for legged robots," *Science Robotics*, vol. 4, no. 26, p. eaau5872, 2019.
- [9] I. Radosavovic, T. Xiao, B. Zhang, T. Darrell, J. Malik, and K. Sreenath, "Learning humanoid locomotion with transformers," *arXiv preprint arXiv:2303.03381*, 2023.
- [10] B. Weng, G. A. Castillo, W. Zhang, and A. Hereid, "On the comparability and optimal aggressiveness of the adversarial scenario-based safety testing of robots," *IEEE Transactions on Robotics*, vol. 39, pp. 3299–3318, 2023.
- [11] —, "On safety testing, validation, and characterization with scenario-sampling: A case study of legged robots," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 5179–5186.
- [12] F. Aller, D. Pinto-Fernandez, D. Torricelli, J. L. Pons, and K. Mombaur, "From the state of the art of assessment metrics toward novel concepts for humanoid robot locomotion benchmarking," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 914–920, 2019.
- [13] DubstepDinosaurs, "Atlas the next generation robot - harassed with hockey stick!" <https://www.youtube.com/watch?v=E0Rc9CzVRuQ>, Feb 2016.
- [14] E. S. Walsh, P. Rousseau, and T. B. Hoshizaki, "The influence of impact location and angle on the dynamic impact response of a hybrid iii headform," *Sports Engineering*, vol. 13, pp. 135–143, 2011.
- [15] M. A. Allison, Y. S. Kang, M. R. Maltese, J. H. Bolte, and K. B. Arbogast, "Measurement of hybrid iii head impact kinematics using an accelerometer and gyroscope system in ice hockey helmets," *Annals of biomedical engineering*, vol. 43, pp. 1896–1906, 2015.
- [16] Y.-S. Kang, A. Bendig, J. Stammen, E. Hutter, K. Moorhouse, J. H. Bolte IV, and A. M. Agnew, "Comparison of small female pmhs thoracic responses to scaled response corridors in a frontal hub impact," *Traffic Injury Prevention*, vol. 24, no. 1, pp. 62–68, 2023.
- [17] E. D. Edwards, T. Landry, M. Jesunathadas, T. A. Plaisted, R. J. Neice, T. E. Gould, M. Kleinberger, and S. G. Piland, "Kinematic assessment of the nocsae headform during blunt impacts with a pneumatic linear impactor," *Sports Engineering*, vol. 26, no. 1, p. 13, 2023.
- [18] M. A. Meyers and K. K. Chawla, *Mechanical behavior of materials*. Cambridge university press, 2008.
- [19] J. Varghese and A. Dasgupta, "Test methodology for impact testing of portable electronic products," in *ASME International Mechanical Engineering Congress and Exposition*, vol. 37149, 2003, pp. 71–77.
- [20] D. C. Viano and D. Halstead, "Change in size and impact performance of football helmets from the 1970s to 2010," *Annals of biomedical engineering*, vol. 40, pp. 175–184, 2012.
- [21] L. V. Burgin and R. M. Aspden, "A drop tower for controlled impact testing of biological tissues," *Medical engineering & physics*, vol. 29, no. 4, pp. 525–530, 2007.
- [22] M. Cheng, J.-P. Dionne, and A. Makris, "On drop-tower test methodology for blast mitigation seat evaluation," *International Journal of Impact Engineering*, vol. 37, no. 12, pp. 1180–1187, 2010.
- [23] W. Li, S. Li, L. Yan, D. Qin, S. Sun, and H. Zhao, "Development and field application of a pulse-jet hydraulic impactor," *Natural Gas Industry B*, vol. 5, no. 6, pp. 558–564, 2018.
- [24] C. H. Mejia, J. Jayet, P. Germano, A. Thabuis, and Y. Perriard, "Linear impact generator for automated dataset acquisition of elastic waves in haptic surfaces," in *2019 22nd International Conference on Electrical Machines and Systems (ICEMS)*. IEEE, 2019, pp. 1–5.
- [25] S. J. Rao, G. J. Forkenbrock *et al.*, "Test procedures traffic jam assist test development considerations." United States. Department of Transportation. National Highway Traffic Safety Administration, Tech. Rep., 2019.
- [26] G. A. Castillo, B. Weng, S. Yang, W. Zhang, and A. Hereid, "Template model inspired task space learning for robust bipedal locomotion," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023, pp. 8582–8589.
- [27] A. Del Prete, N. Mansard, O. E. Ramos, O. Stasse, and F. Nori, "Implementing torque control with high-ratio gear boxes and without joint-torque sensors," *International Journal of Humanoid Robotics*, vol. 13, no. 01, p. 1550044, 2016.
- [28] M. H. Raibert, J. Brown, C. H. B., K. Michael, and H. Jeff, "Dynamically stable legged locomotion," 1989. [Online]. Available: <https://dspace.mit.edu/handle/1721.1/6820>
- [29] J. Lee, J. Ahn, D. Kim, S. H. Bang, and L. Sentis, "Online gain adaptation of whole-body control for legged robots with unknown disturbances," *Frontiers in Robotics and AI*, vol. 8, p. 788902, 2022.
- [30] G. Romualdi, S. Dafarra, G. L'Erario, I. Sorrentino, S. Traversaro, and D. Pucci, "Online non-linear centroidal mpc for humanoid robot locomotion with step adjustment," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 10 412–10 419.
- [31] T. J. Walilko, D. C. Viano, and C. A. Bir, "Biomechanics of the head for olympic boxer punches to the face," *British journal of sports medicine*, vol. 39, no. 10, pp. 710–719, 2005.
- [32] P. Akella and A. D. Ames, "A barrier-based scenario approach to verifying safety-critical systems," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 11 062–11 069, 2022.