

Custom Sine Waves Are Enough for Imitation Learning of Bipedal Gaits with Different Styles

Authors: Qi Wu^{*1}, Chong Zhang^{*2}, Yanchen Liu³

¹Department of Mechanical Engineering, Tsinghua University

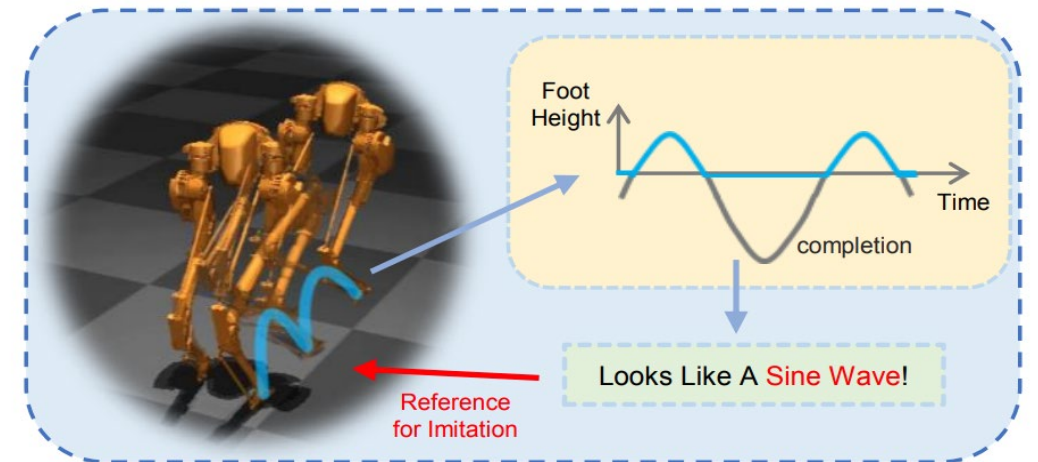
²Department of Precision Instrument, Tsinghua University

³Department of Informatics, Technical University of Munich

^{*}Equal contribution

Corresponding: Qi Wu, wuqi19@mails.tsinghua.edu.cn

Presenter: Chong Zhang

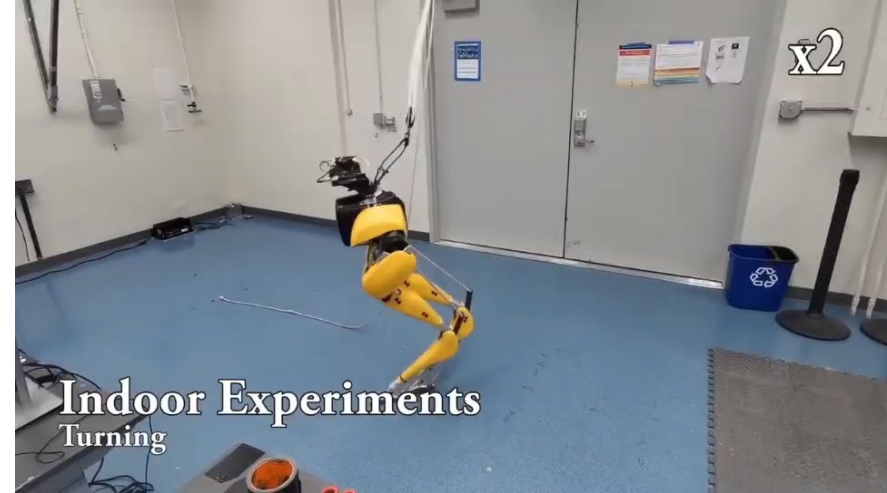


Introduction

- Deep Reinforcement Learning (DRL)
 - A new solution to legged locomotion
 - An easy solution to dynamics modelling



[Miki, T. , et al. Science Robotics, 2022]

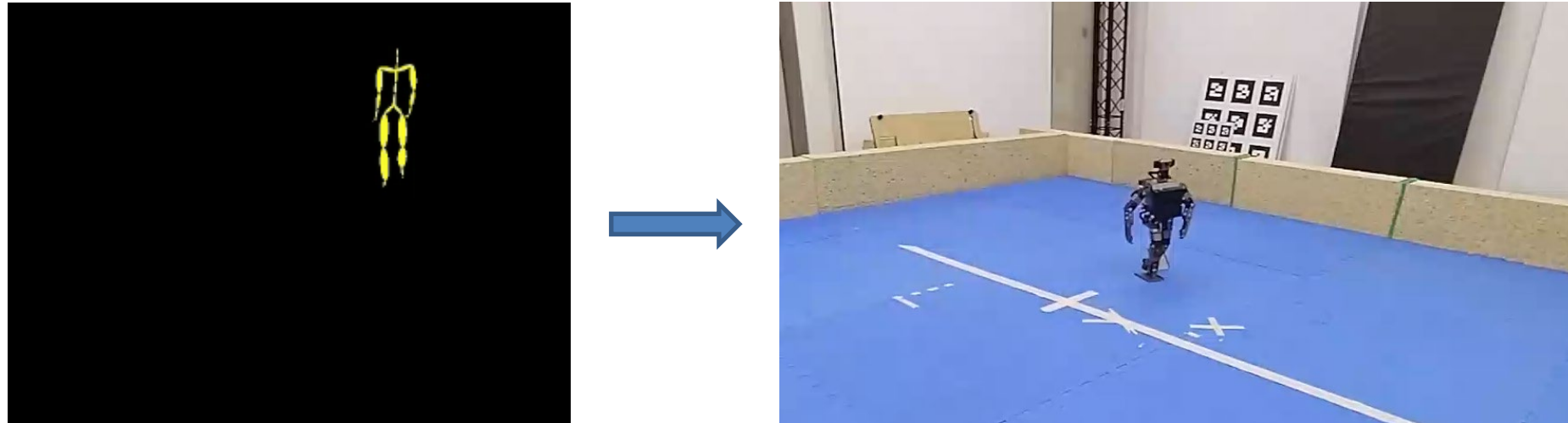


[Li, Z. , et al. ICRA 2021]

Introduction

- DRL implementation for bipedal gaits
 - End-2-end learning: time-consuming
 - Dozens of hours for training [Siekman, J., et al. RSS 2021, & ICRA2021]
 - GPU-based simulation hasn't shown enough fidelity for bipedal robots.
 - Motion imitation: costly mocaps or manually tuned references
 - Human motion capture is costly and anthropomorphic.

Human motion capture: costly, limited styles, anthropomorphic!



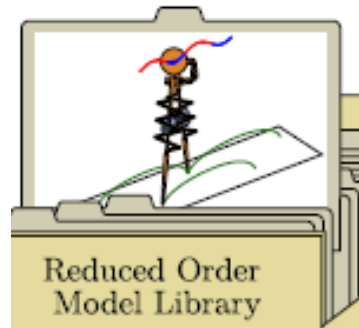
Bohez, Steven, et al. Imitate and Repurpose: Learning Reusable Robot Movement Skills From Human and Animal Behaviors.

Introduction

- DRL implementation for bipedal gaits
 - End-2-end learning: time-consuming
 - Dozens of hours for training [Siekman, J., et al. RSS 2021, & ICRA2021]
 - GPU-based simulation hasn't shown enough fidelity for bipedal robots.
 - Motion imitation: costly mocaps or manually tuned references
 - Human motion capture is costly and anthropomorphic.
 - Manually tuned references require lots of expert knowledge and efforts.

Manually tuned reference: cheaper than mocaps, requiring knowledge and efforts

Spring Mass Model



Green, Kevin, et al. "Learning spring mass locomotion: Guiding policies with a reduced-order model."

Other choices:

Manually tuned controller: Xie, Z. , et al. IROS 2018

HZD gait library: Li, Z. , et al. ICRA 2021

...

Introduction

- DRL implementation for bipedal gaits
 - End-2-end learning: time-consuming
 - Dozens of hours for training [Siekman, J., et al. RSS 2021, & ICRA2021]
 - GPU-based simulation hasn't shown enough fidelity for bipedal robots.
 - Motion imitation: costly mocaps or manually tuned references
 - Human motion capture is costly and anthropomorphic.
 - Manually tuned references require lots of expert knowledge and efforts.
 - Gait styles are limited to the references and difficult to tune.
 - Massive expert efforts for reward & hyperparameter tuning

Introduction

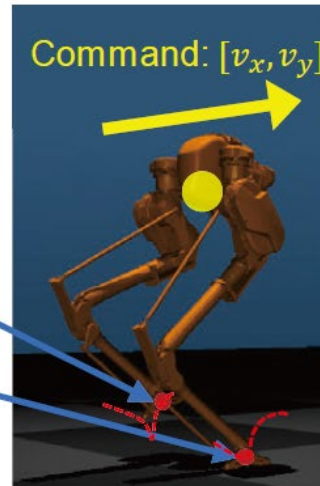
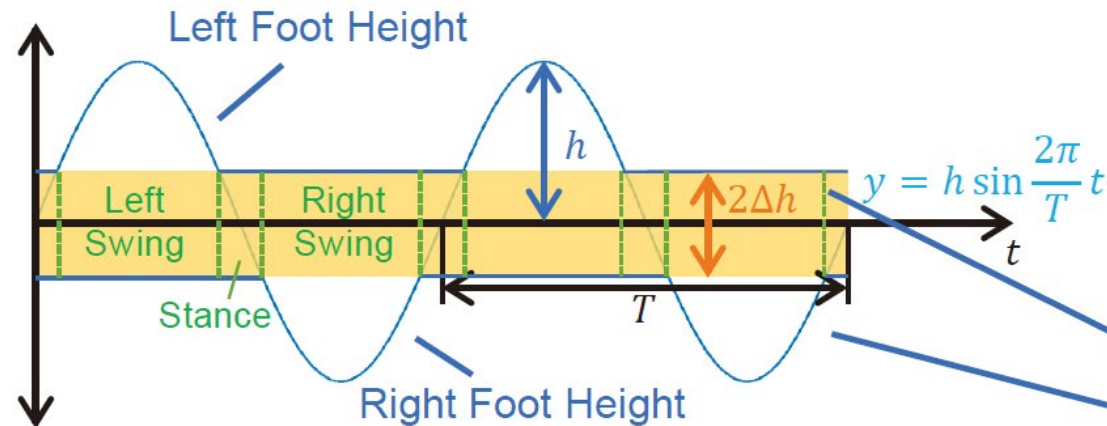
- Our motivation
 - Can the references be cheap and easy to obtain?
 - A sine wave suffices!
 - Can the rewards be easy to tune?
 - Few reward terms without special efforts to tune!
 - Can the learning process be efficient?
 - As efficient as those with manually-tuned references!
 - Anything else?
 - Deploy a very simple system and focus on skill learning, as suggested by previous works [Siekman, J., et al. RSS 2021]:
sim2real and skill learning can be separately treated.

Our Method

- The very simple reference
 - Easy to generate
 - Easy to tune the styles

- The walking styles

- $h - \Delta h$ for foot clearance: higher or lower
- $\frac{\Delta h}{h}$ for double-support span: also faster or slower feet up and down
- T for frequency: higher or lower

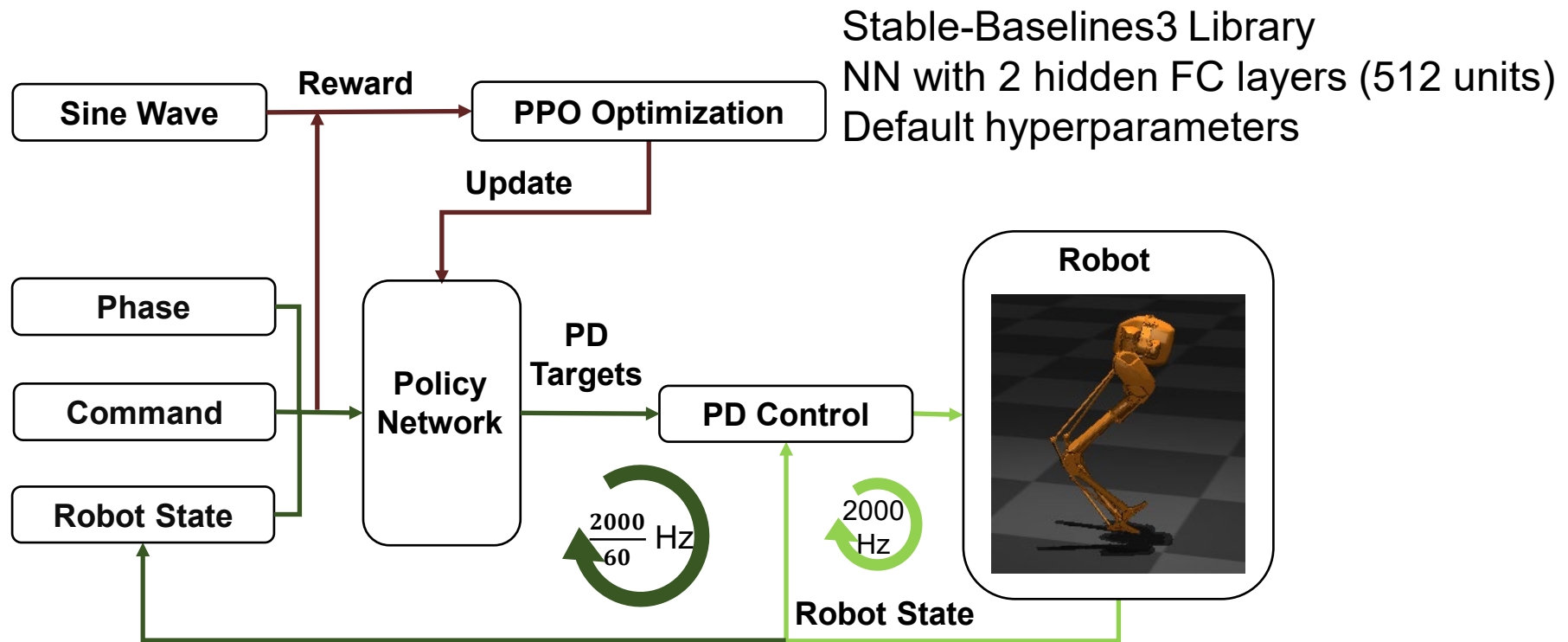


$$h_{\text{ref}_1} = h_{\text{left}} = \max(0, h \sin(\frac{2\pi}{T}t + \phi_0) - \Delta h),$$

$$h_{\text{ref}_2} = h_{\text{right}} = \max(0, h \sin(\frac{2\pi}{T}t + \phi_0 + \pi) - \Delta h).$$

Our Method

- The very simple system



Our Method

- The very simple rewards

[Peng, X. B. , et al. SIGGRAPH 2018]: joint position, joint velocity, end-effector position, and CoM position (4) terms for imitation; task specific terms for performance

[Xie, Z. , et al. IROS 2018]: joint position, pelvis position, and pelvis orientation (3) terms for imitation; shin spring term for regularization

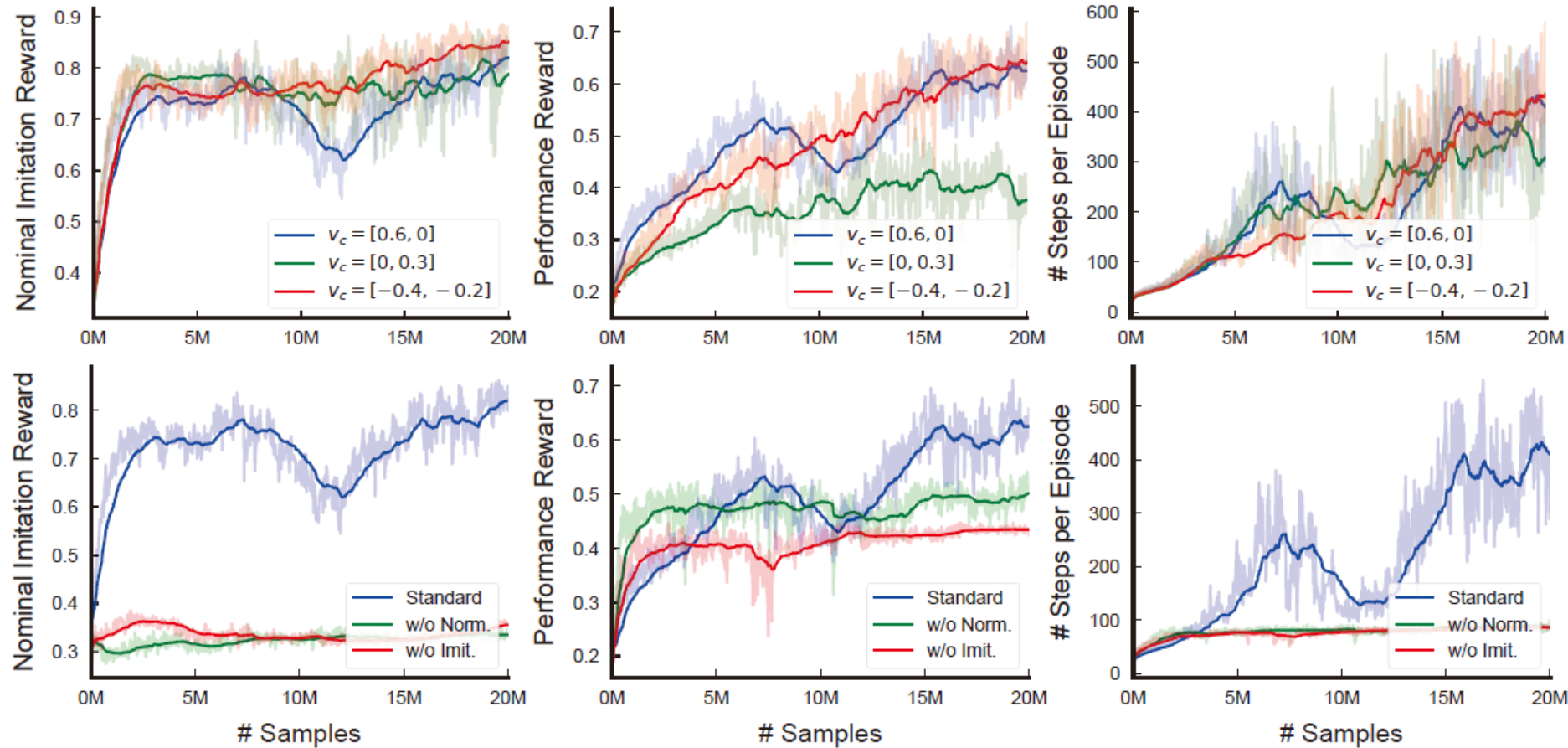
[Li, Z. , et al. ICRA 2021]: joint position, pelvis position, pelvis translational velocity, pelvis rotation, and pelvis rotational velocity (5) terms for imitation; 2 terms for regularization

Ours: 1 term for imitation, 2 terms for performance, 1 (but can be more) term for regularization and 1 term for early termination

Our Method

- The very simple rewards: $0.5r^I + 0.5r^P + r^R + r^T$
 - Imitation:
 - $r^{I*} = \exp(-\frac{\text{dist}^2(h_{\text{ref}}, h_{\text{foot}})}{0.05^2})$, $r^I = \frac{r^{I*} - 0.4}{1.0 - 0.4}$: a penalty for surviving without pursuing imitation
 - Performance:
 - $r^P = 0.75 \exp\left(-\frac{\text{dist}^2(v_{\text{pelvis}}, v_{\text{command}})}{\max(0.1^2, 0.5\|v_{\text{command}}\|^2)}\right) + 0.25 \exp\left(-\frac{\text{dist}^2(ori_{\text{pelvis}}, ori_{\text{standing}})}{0.1}\right)$
 - Regularization:
 - $r^R = 0.1 \exp(-\frac{\text{dist}^2(q_{\text{shin}}, \mathbf{0})}{0.001})$
 - Termination:
 - $r^T = -10$ if early termination else 0: triggered when fall or “derail” too far

Results



3-5 hours on an ordinary desktop!

Existing works:

Dozens of M steps for imitation learning

Hundreds of M steps for end2end learning

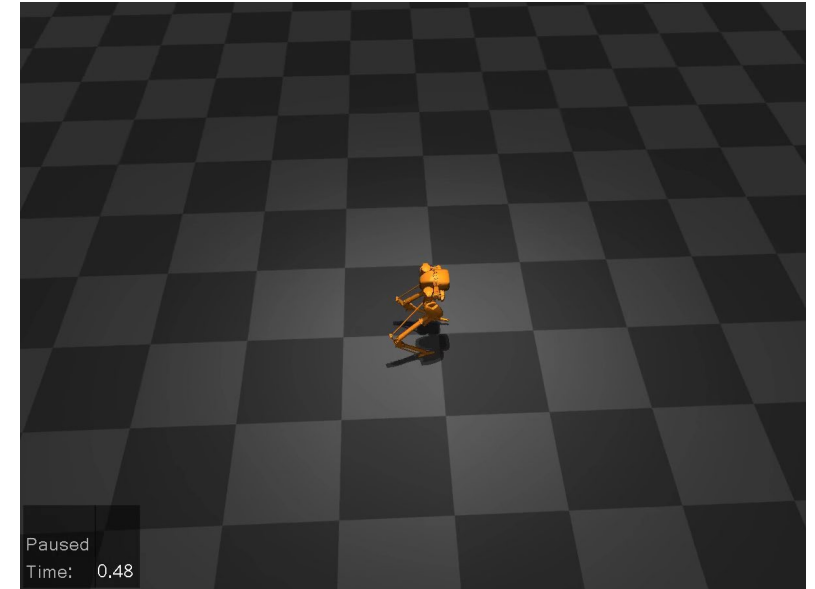
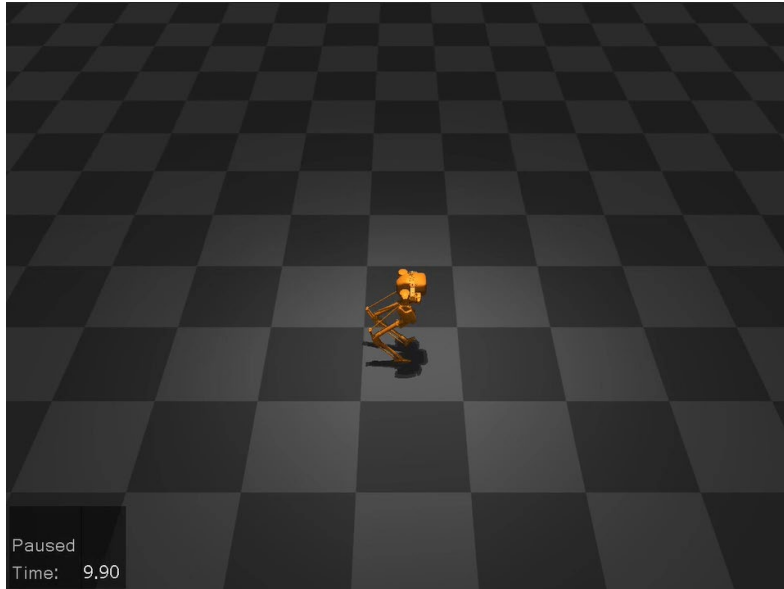
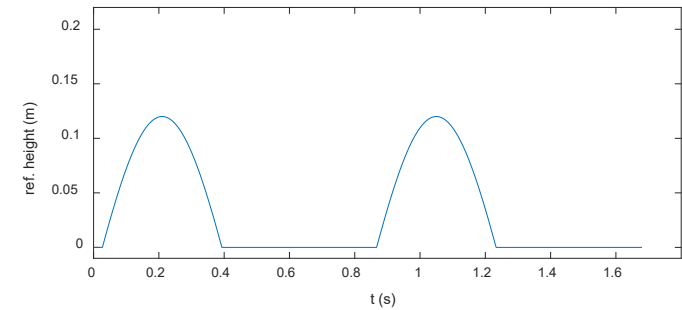
Results

- Same reference, different velocities

$$v_x = 0.6\text{m/s}, v_y = 0\text{m/s}$$

$$v_x = 0\text{ m/s}, v_y = 0.3\text{m/s}$$

$$v_x = -0.4\text{m/s}, v_y = -0.2\text{m/s}$$

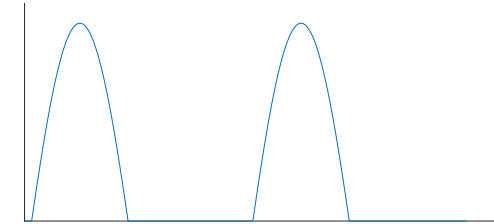
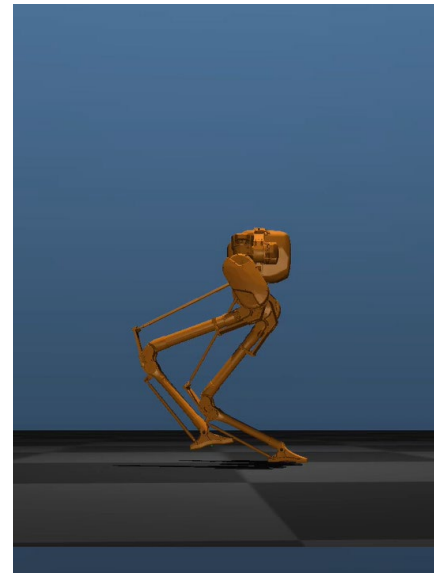
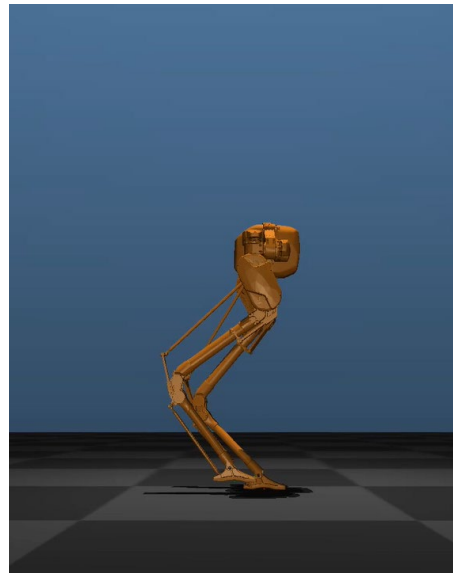
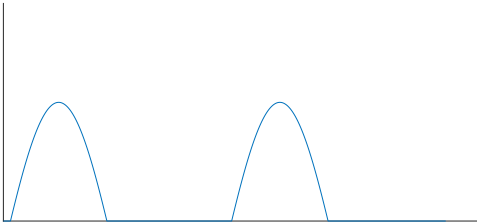
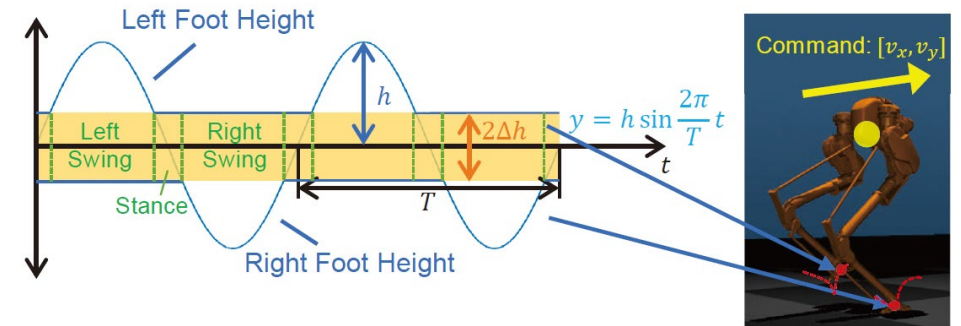


Results

- Different styles: changing $h - \Delta h$

$$h - \Delta h = 0.12$$

$$h - \Delta h = 0.2$$

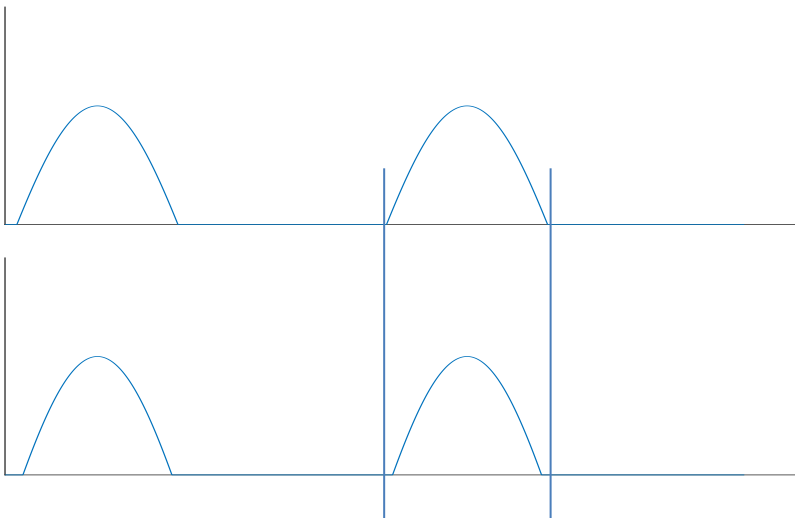


Results

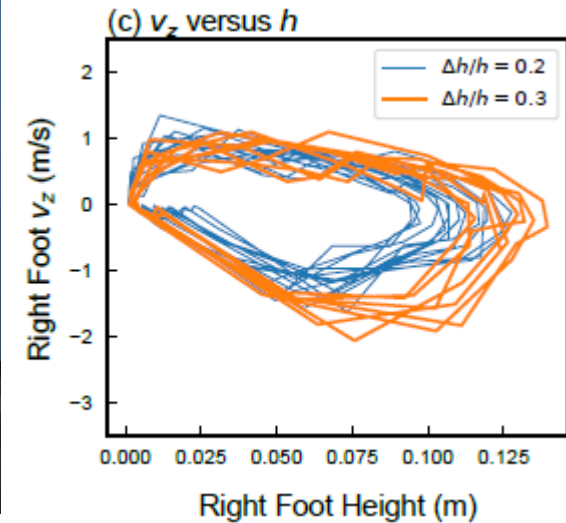
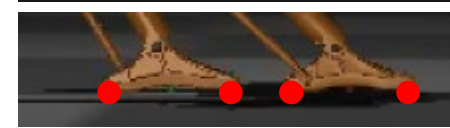
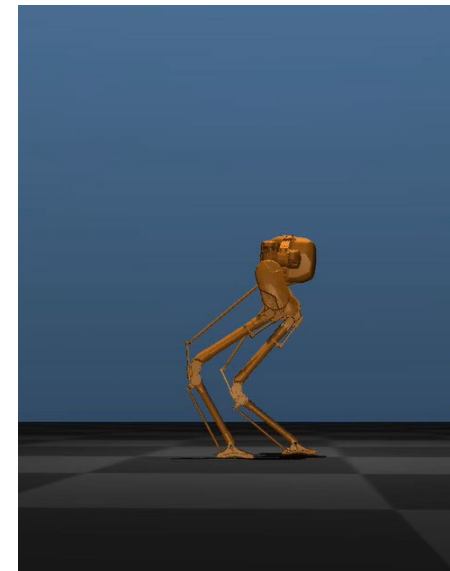
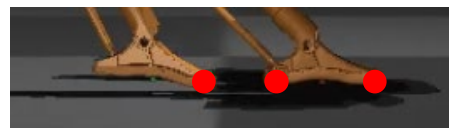
- Different styles: changing $\Delta h/h$

$$\Delta h/h = 0.2$$

$$\Delta h/h = 0.3$$



Contact Points Before Lift

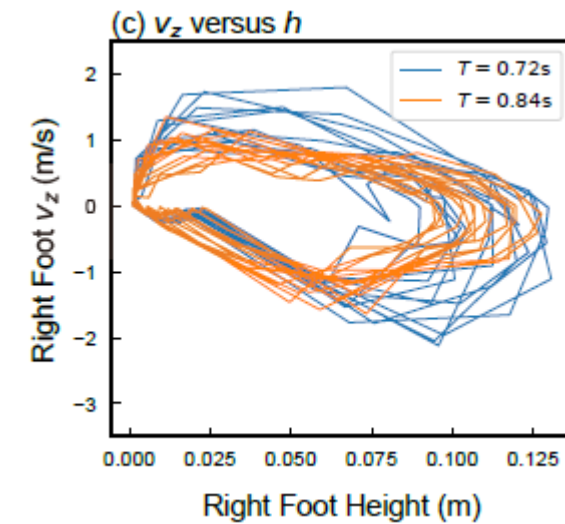
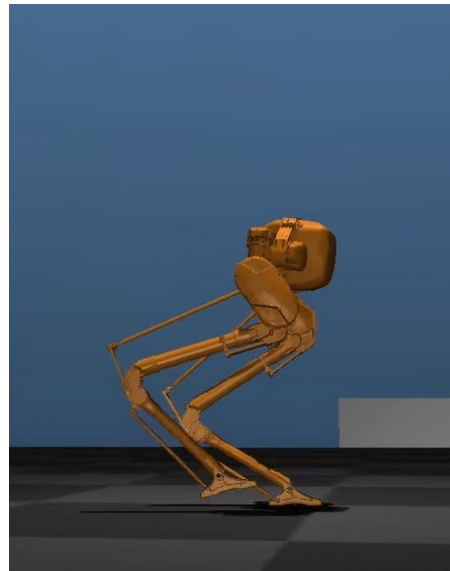
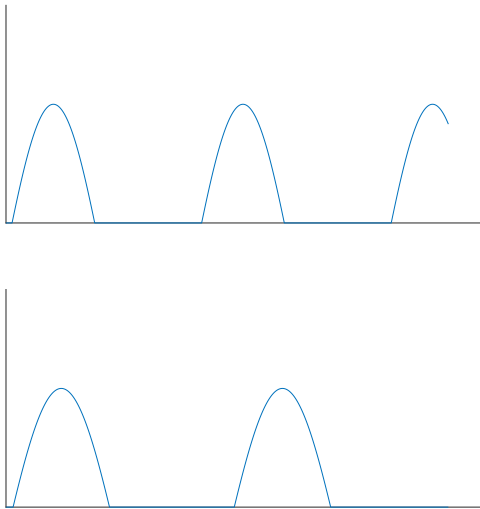


Results

- Different styles: changing T

$T = 0.72 \text{ s}$

$T = 0.84 \text{ s}$



Future Works

- What we have done:
 - Using this technique to verify the agility of a bipedal robot that is undergoing iterative design
- What's the next:
 - To adapt the sine waves to more periodical motions
 - To achieve sim2real on different robot platforms

Thanks for watching!

Qi Wu^{*1}, Chong Zhang^{*2}, Yanchen Liu³

¹Department of Mechanical Engineering, Tsinghua University

²Department of Precision Instrument, Tsinghua University

³Department of Informatics, Technical University of Munich

Corresponding: Qi Wu, wuqi19@mails.tsinghua.edu.cn

Codes and this PPT: <https://github.com/WooQi57/sin-cassie-rl>

