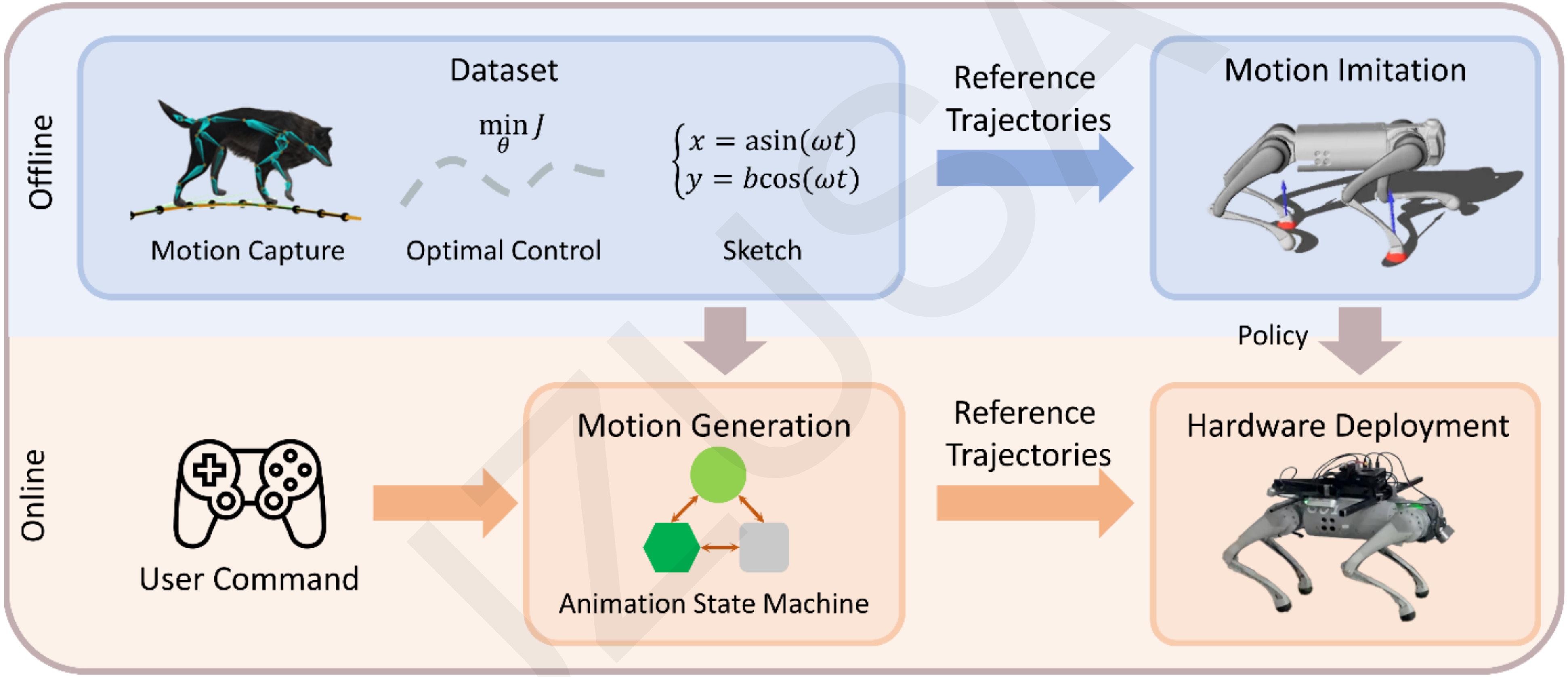
A learning-based control pipeline for generic motor skills for quadruped robots

Yecheng SHAO, Yongbin JIN, Zhilong HUANG, Hongtao WANG, Wei YANG

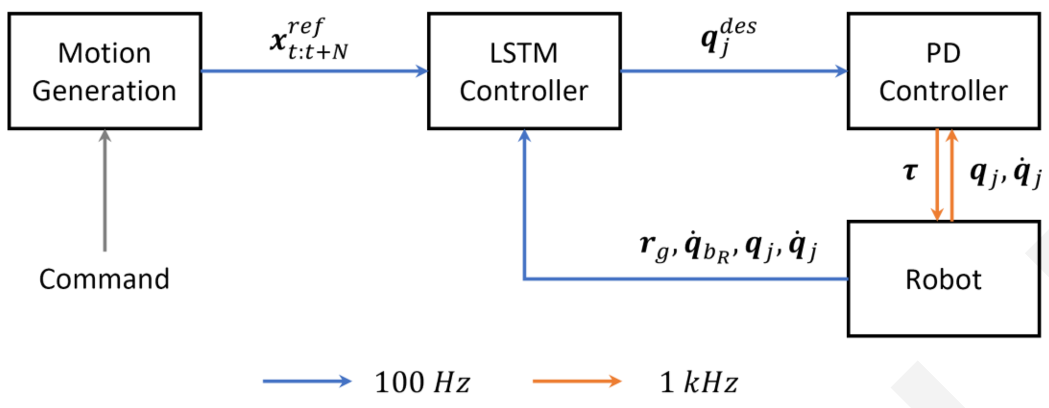
Cite this as: Yecheng SHAO, Yongbin JIN, Zhilong HUANG, Hongtao WANG, Wei YANG, 2024. A learning-based control pipeline for generic motor skills for quadruped robots. *Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering)*, 25(6):443-454. https://doi.org/10.1631/jzus.A2300128

Overall Control Pipeline



- **Offline Training:** Train the policy to imitate motions from presented dataset.
- **Online Control:** Use state machine to generate motions according to user command based on the dataset and use the trained policy to track the generated motions.

Reinforcement Learning

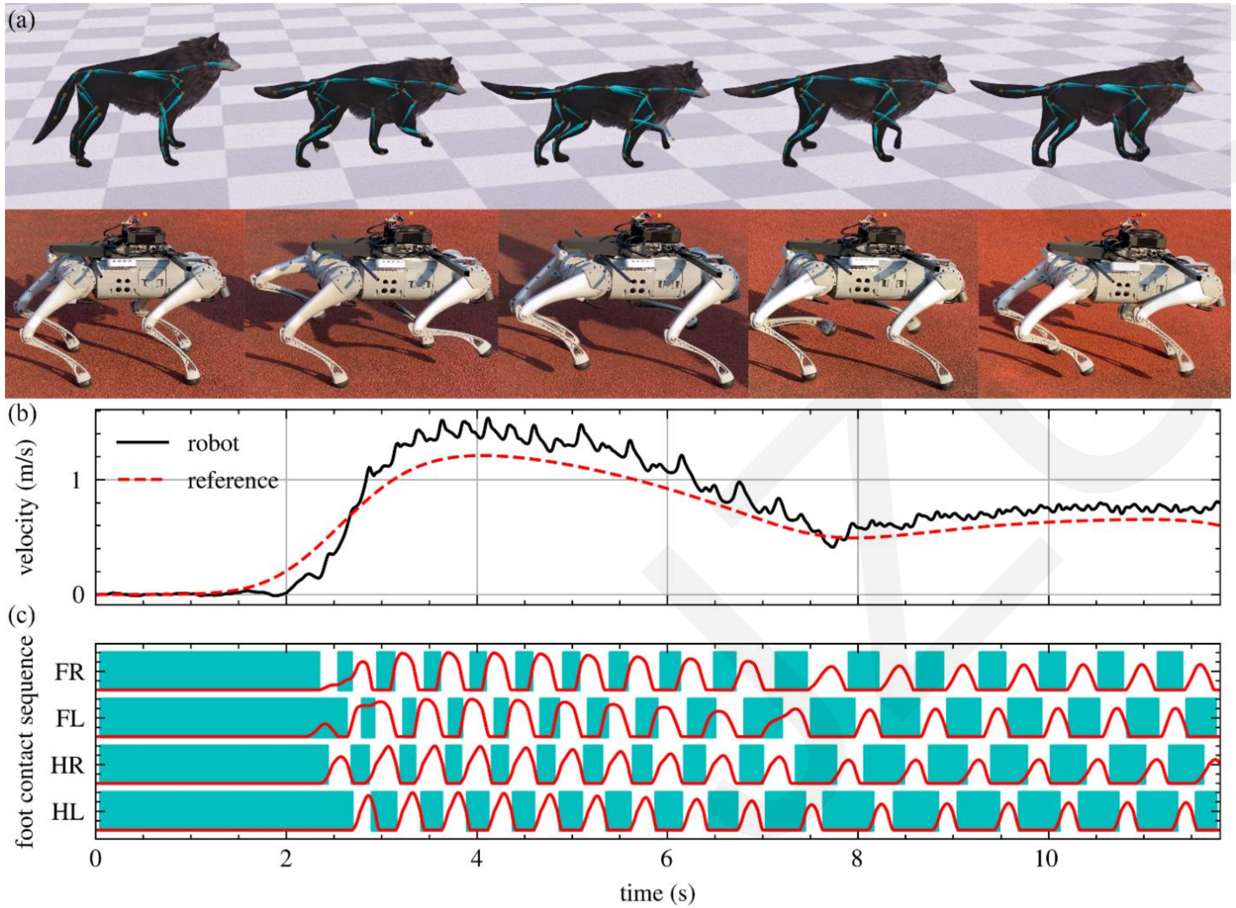


Reward	Weight
Base Velocity Tracking	0.2
Base Height Tracking	0.1
Base Orientation Tracking	0.1
Torque Penalty	0.2
Joint Position/Velocity Tracking	0.4

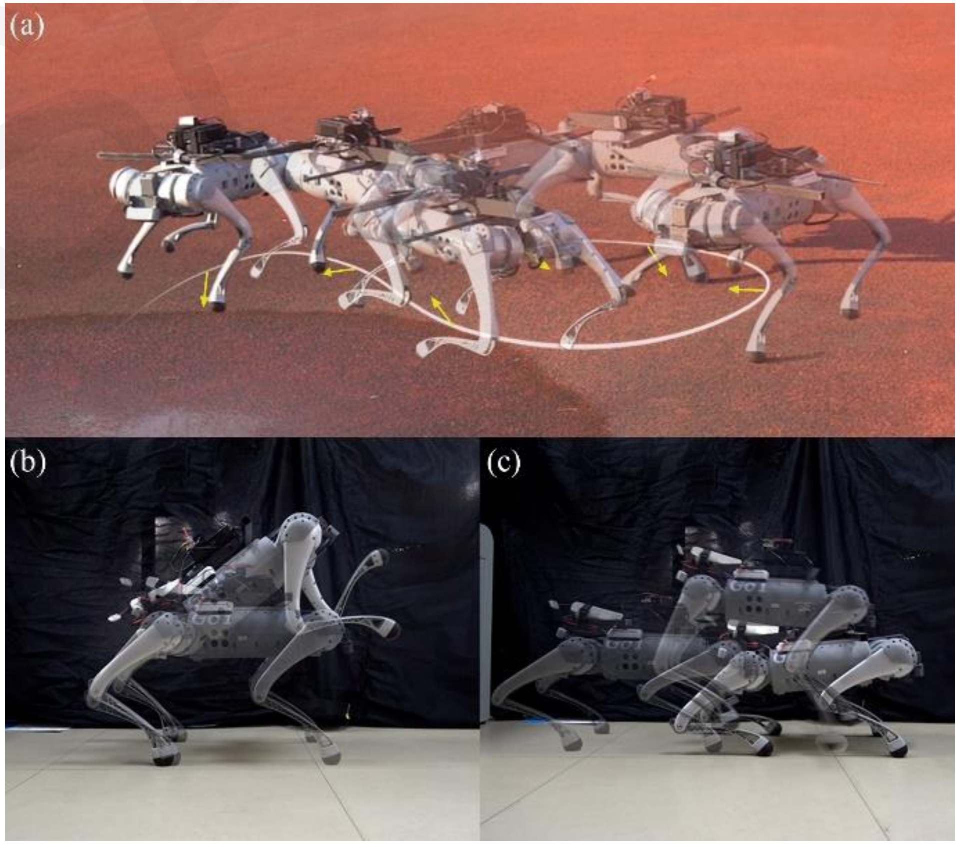
	Notion	Dimension	Description
Observations	q_j	12	Joint positions.
	$\dot{q}_j$	12	Joint velocities.
	$\dot{q}_{b_R}$	3	Angular velocities of the trunk.
	r_g	3	Angular positions of the trunk.
	$x_{t:t+N}^{ref}$	$19N$	N steps of future reference trajectories.
Actions	q_j^{des}	12	The desired joint position.

Various Motions from the Proposed Methods

Natural Gait Transitions



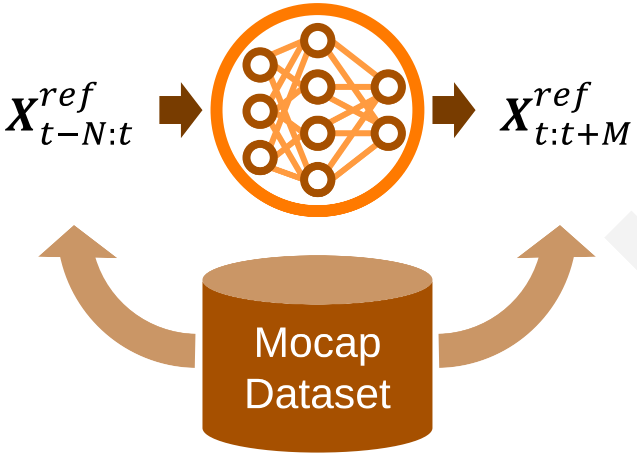
Commanded Locomotion



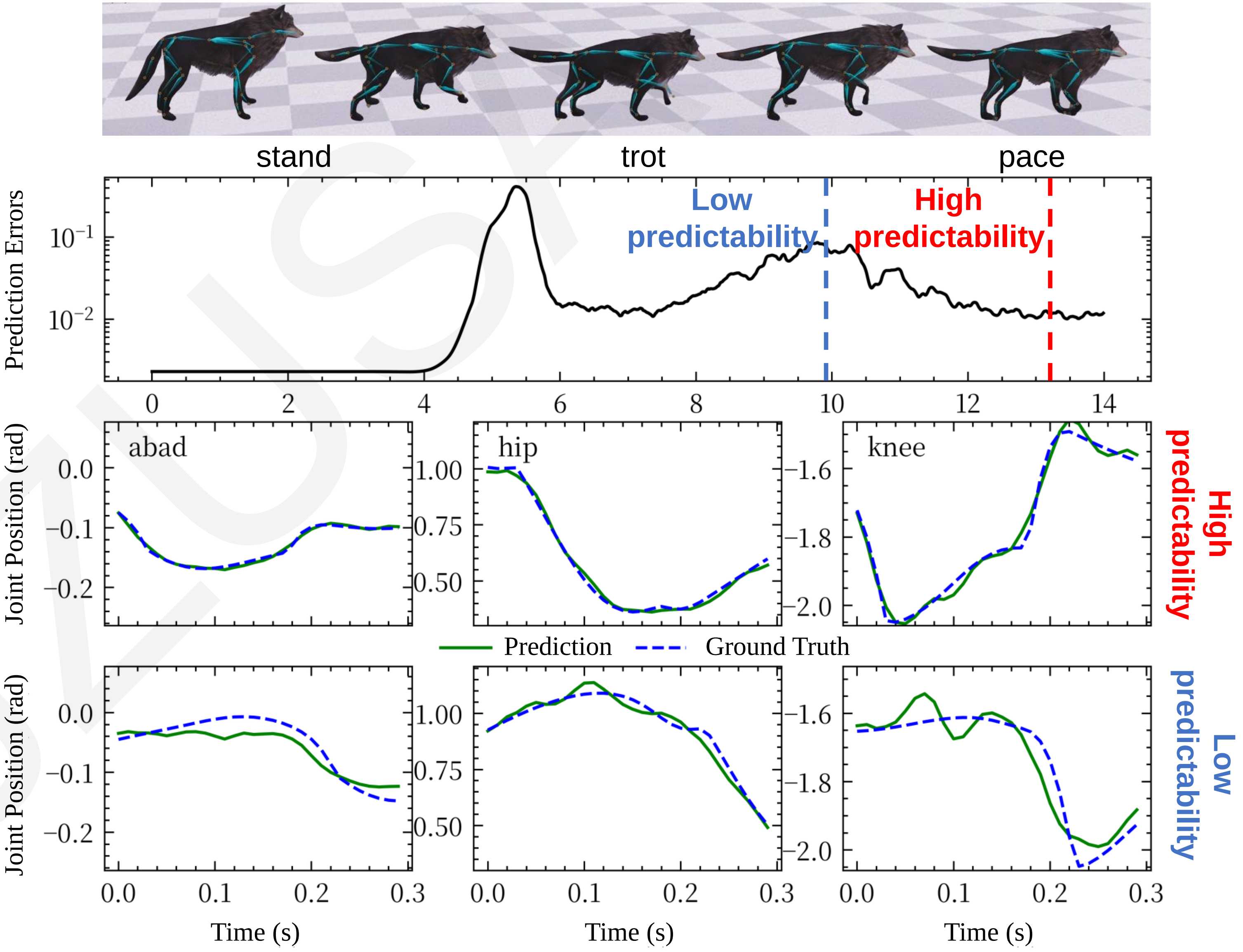
The Correlation Inside the Reference Trajectories

The necessary length of the reference trajectories is closely related to the internal correlation of the references.

We fit the reference trajectories with a neural network and use the prediction error as an indicator to the correlation.

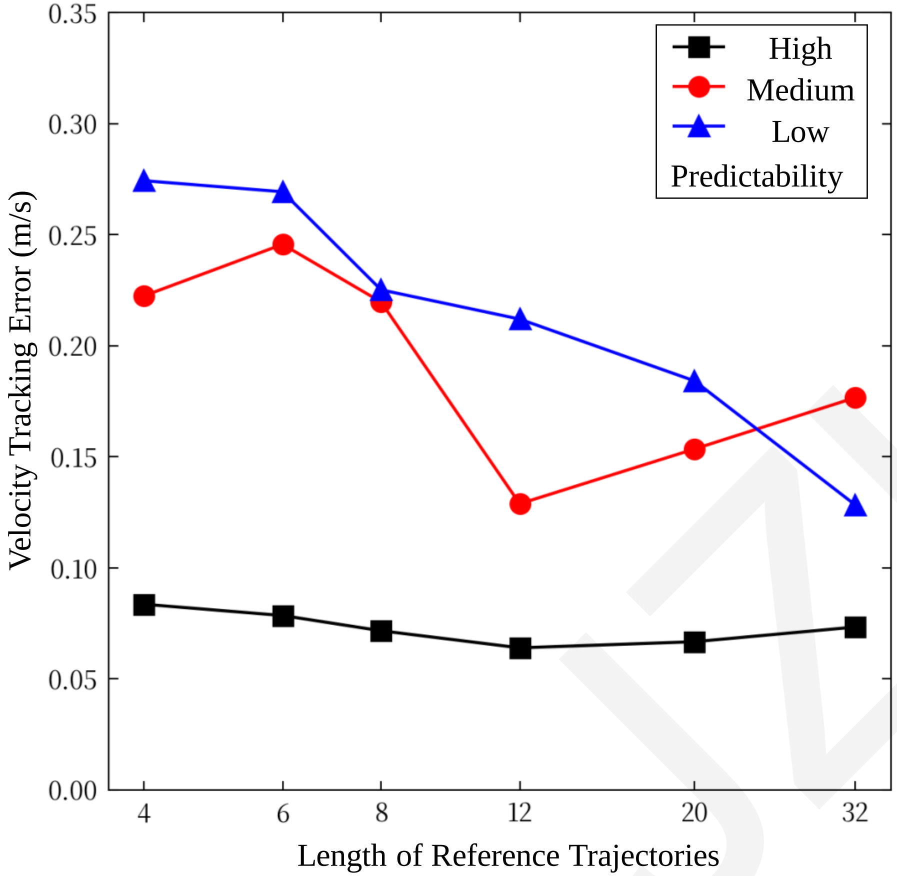


$$\text{Prediction Error}(t) = \left\| X^{ref}_{t:t+M} - \hat{X}^{ref}_{t:t+M} \right\|$$

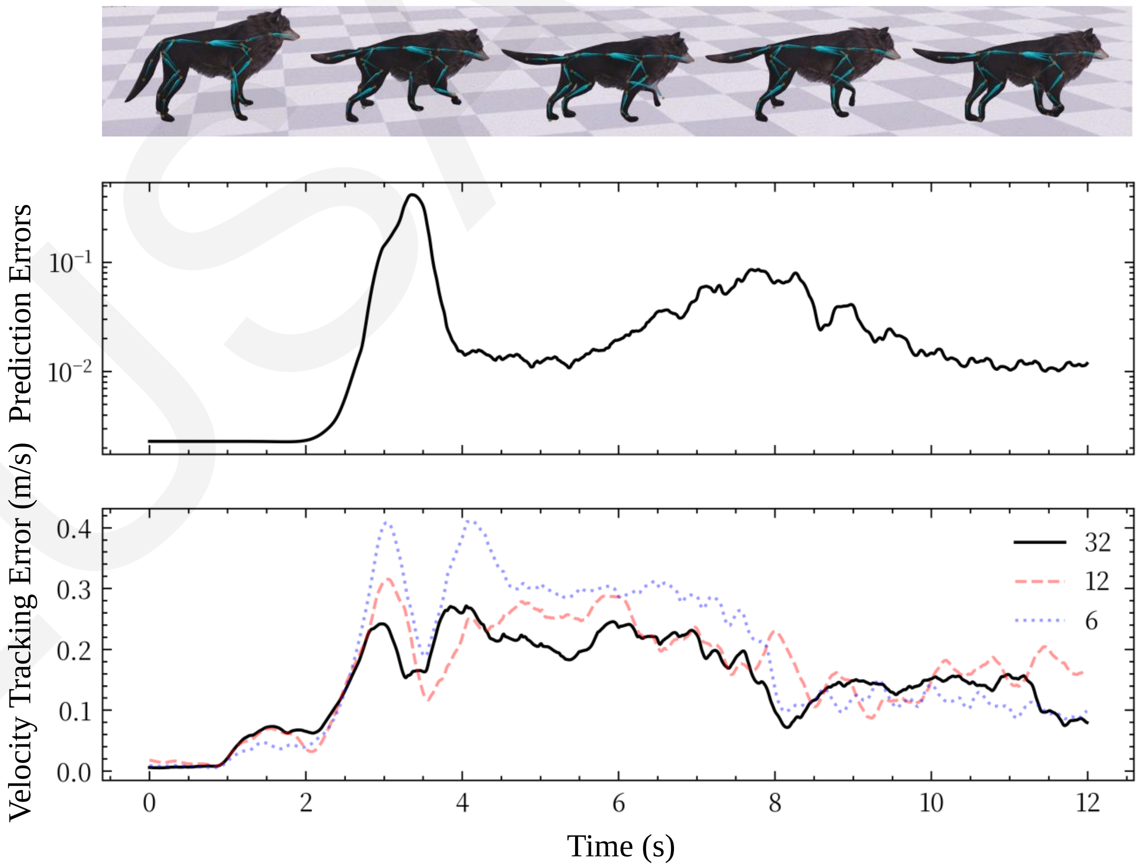


The Effect on the Control Performance

Statistics



Time Series



- When the lengths of reference trajectories are same, higher predictability will bring higher performance.
- For motions with low predictability, longer reference trajectories can improve the control performance.

Conclusions

- We introduce a control pipeline combining motion synthesis and motion imitation-based RL for generic motor skills.
- A single policy is able to learn various motor skills simultaneously under the proposed control pipeline.
- The trained policy can uncover the correlations lurking behind the reference motions to improve control performance.