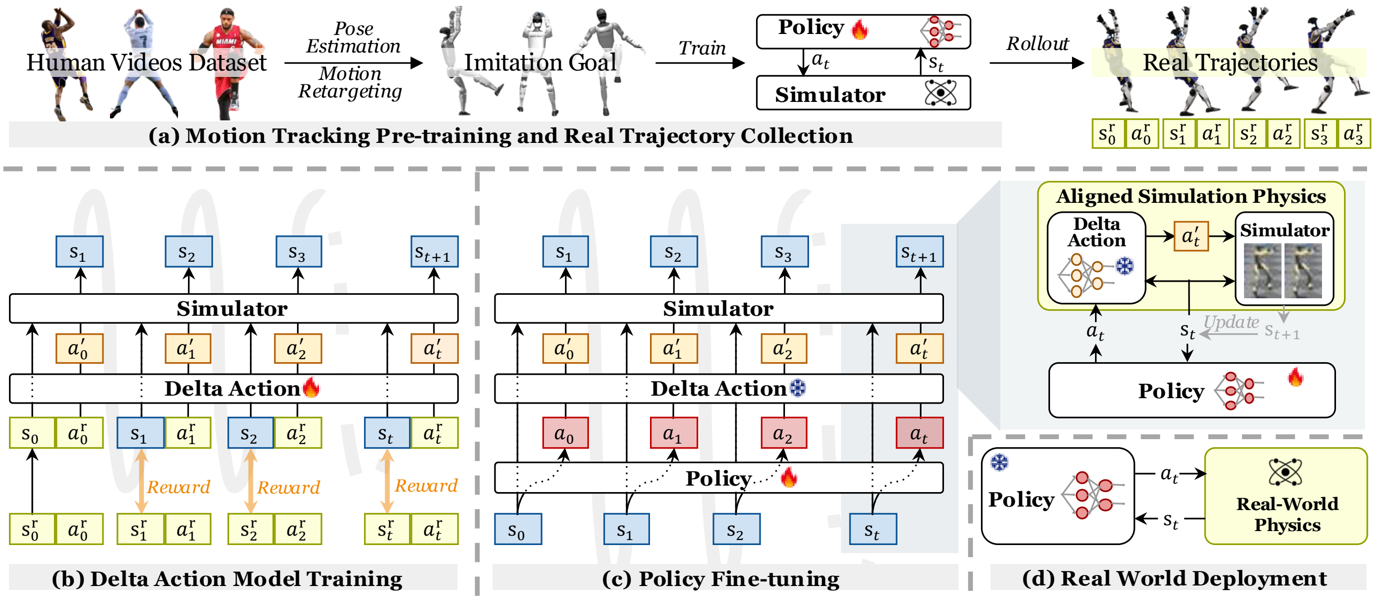

Fig. 2. Overview of ASAP. (a) **Motion Tracking Pre-training and Real Trajectory Collection**: With the humanoid motions retargeted from human videos, we pre-train multiple motion tracking policies to roll out real-world trajectories. (b) **Delta Action Model Training**: Based on the real-world rollout data, we train the delta action model by minimizing the discrepancy between simulation state s_t and real-world state s_t^r. (c) **Policy Fine-tuning**: We freeze the delta action model, incorporate it into the simulator to align the real-world physics and then fine-tune the pre-trained motion tracking policy. (d) **Real-World Deployment**: Finally, we deploy the fine-tuned policy directly in the real world without delta action model.

to train a delta (residual) action model that compensates for the dynamics mismatch. Then ASAP fine-tunes pre-trained policies with the delta action model integrated into the simulator to align effectively with real-world dynamics. We evaluate ASAP across three transfer scenarios—IsaacGym to IsaacSim, IsaacGym to Genesis, and IsaacGym to the real-world Unitree G1 humanoid robot. Our approach significantly improves agility and whole-body coordination across various dynamic motions, reducing tracking error compared to SysID, DR, and delta dynamics learning baselines. ASAP enables highly agile motions that were previously difficult to achieve, demonstrating the potential of delta action learning in bridging simulation and real-world dynamics. These results suggest a promising sim-to-real direction for developing more expressive and agile humanoids.

I. Introduction

For decades, we have envisioned humanoid robots achieving or even surpassing human-level agility. However, most prior work [46, 74, 47, 73, 107, 19, 95, 50] has primarily focused on locomotion, treating the legs as a means of mobility. Recent studies [10, 25, 24, 26, 32] have introduced whole-body expressiveness in humanoid robots, but these efforts have primarily focused on upper-body motions and have yet to achieve the agility seen in human movement. Achieving agile, whole-body skills in humanoid robots remains a fundamental challenge due to not only hardware limits but also the mismatch between simulated dynamics and real-world physics.

Three main approaches have emerged to bridge the dynamics mismatch: System Identification (SysID) methods, domain randomization (DR), and learned dynamics methods. SysID methods directly estimate critical physical parameters, such as motor response characteristics, the mass of each robot link, and terrain properties [102, 19]. However, these methods require a pre-defined parameter space [49], which may not fully capture the sim-to-real gap, especially when real-world dynamics fall outside the modeled distribution. SysID also often relies on ground truth torque measurements [29], which are unavailable on many widely used hardware platforms, limiting its practical applicability. DR methods, in contrast, first train control policies in simulation before deploying them on real-world hardwares [85, 79, 59]. To mitigate the dynamics mismatch between simulation and real-world physics, DR methods rely on randomizing simulation parameters [87, 68]; but this can lead to overly conservative policies [25], ultimately hindering the development of highly agile skills. Another approach to bridge dynamics mismatch is learning a dynamics model of real-world physics using real-world data. While this approach has demonstrated success in low-dimensional systems such as drones [81] and ground vehicles [97], its effectiveness for humanoid robots remains unexplored.

To this end, we propose ASAP, a two-stage framework that aligns the dynamics mismatch between simulation and real-world physics, enabling agile humanoid whole-body skills. ASAP involves a pre-training stage where we train base policies in simulation and a post-training stage that finetunes the policy by aligning simulation and real-world dynamics. In the **pre-training** stage, we train a motion tracking policy in simulation using human motion videos as data sources. These motions are first retargeted to humanoid robots [25], and a phase-conditioned motion tracking policy [67] is trained to follow the retargeted movements. However, directly deploying this policy on real hardware results in degraded performance due to the dynamics mismatch. To address this, the **post-training** stage collects real-world rollout data, including proprioceptive states and positions recorded by the motion capture