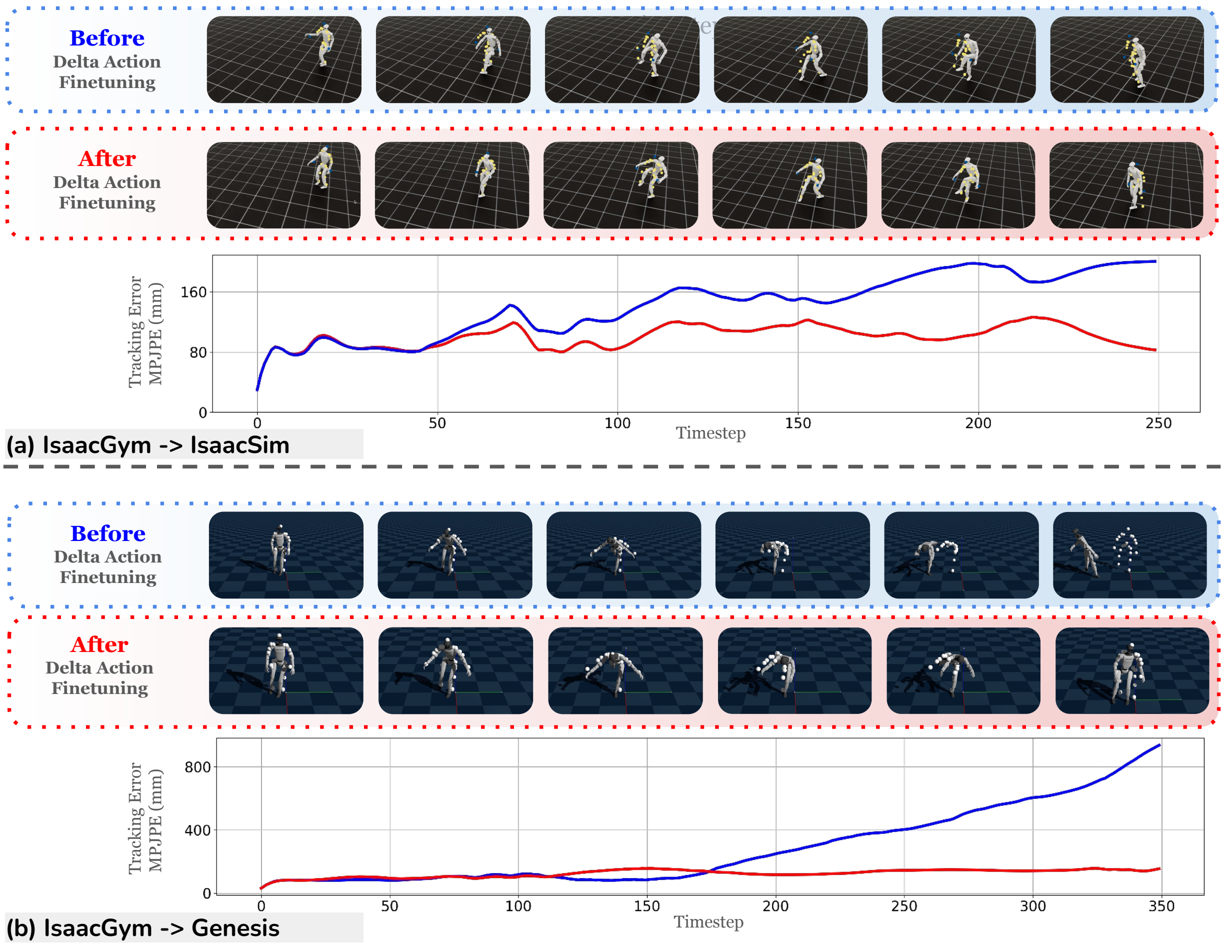

Fig. 7. Visualization of G1 motion tracking before and after ASAP fine-tuning in IsaacGym, IsaacSim and Genesis. Top: LeBron James' "Silencer" motion tracking policy fine-tuning for IsaacGym to IsaacSim. Bottom: *single foot balance* motion tracking policy fine-tuning for IsaacGym to Genesis.

RL policies deployed without fine-tuning. These visualizations demonstrate that ASAP successfully adapts to new dynamics and maintains stable tracking performance, whereas baseline methods accumulate errors over time, leading to degraded tracking capability. These results highlight the robustness and adaptability of our approach in addressing the sim-to-real gap while preventing overfitting and exploitation. The findings validate that ASAP is an effective paradigm for improving closed-loop performance and ensuring reliable deployment in complex real-world scenarios.

C. Real-World Evaluations

To answer **Q3** (*Does ASAP work for sim-to-real transfer?*). We validate ASAP on real-world Unitree G1 robot.

Real-World Data. In the real-world experiments, we prioritize both motion safety and representativeness by selecting five motion-tracking tasks, including (i) *kick*, (ii) *jump forward*, (iii) *step forward and back*, (iv) *single foot balance* and (v) *single foot jump*. However, collecting over 400 real-world motion clips— the minimum required to train the full 23-DoF delta action model in simulation, as discussed in-Section III-B—poses significant challenges. Our experiments involve highly dynamic motions that cause rapid overheating of joint motors, leading to hardware failures (two Unitree G1 robots broke during data collection). Given these constraints, we adopt a more sample-efficient approach by focusing exclusively on learning a 4-DoF ankle delta action model rather than the full-body 23-DoF model. This decision is motivated by two key factors: (1) the limited availability of real-world data makes training the full 23-DoF delta action model infeasible, and (2) the Unitree G1 robot [77] features a mechanical linkage design in the ankle, which introduces a significant sim-to-real gap that is difficult to bridge with conventional modeling techniques [37]. Under this setting, the original 23 DoF delta action model reduces to 4 DoF delta action model, which needs much less data to be trainable. In practice, we collect 100 motion clips, which prove sufficient to train an effective 4-DoF delta action model for real-world scenarios.

We execute the tracking policy 30 times for each task. In addition to these motion-tracking tasks, we also collect 10 minutes of locomotion data. The locomotion policy will be addressed in the next section, which is also utilized to bridge different tracking policies.

Policy Transition. In the real world, we cannot easily reset