We'd like to first thank the reviewers for their constructive feedback. We are grateful for the appreciation of the novelty of the proposed residual force control (RFC) framework and its empirical effectiveness. We also appreciate the suggested references and will include them. Here we aim to address the main questions raised by the reviewers.

(R1) Q1: Motion complexity claims not backed up by results. A: We will tone down the claims w.r.t. DeepMimic in the final version, but we'd like to stress that the improvement over DeepMimic for ballet dance motions are backed by both quantitative (Fig. 2) and qualitative results (video). This is not to take anything away from DeepMimic, but to show that by compensating for the dynamics mismatch with residual forces, one can achieve better motion imitation. We will follow R1's suggestions to put comparison with DeepMimic in context and discuss the differences more carefully.

(R1) Q2: Generalization of RFC and novelty of motion synthesis. A: The human motion synthesis experiments are designed specifically to show the generalization of RFC. Note that during test time the RFC policy will **not** be further trained to imitate the kinematic output, and it is basically tasked to do one-shot imitation. At test time, the motions generated by the kinematic policy are different from training (different data) and as input to the state of the RFC policy they are analogous to the goals in DeepMimic. Without the residual forces, the agent often falls to the ground as demonstrated in the video (please see posing and walking dog motions starting from 3:31). This in fact demonstrates the generalization of RFC since it is more robust to motion variation than w/o RFC. The key contribution of the motion synthesis is not about the cVAE but the message that an RFC policy can one-shot imitate the noisy output of a generative motion prediction model and is able to do so on a large motion dataset (Human3.6M).

(R1) Q3: Can only be applied to simulation. A: Although the focus of this paper is on virtual human motion synthesis, as mentioned by R1 and other reviewers (R2 & R3), the method could be extended to a scaffolding technique for training complex motion policies, which is actually a direction we are already investigating. Another interesting direction is to use the residual forces to quantify the dynamics mismatch, which could be used to inform agent design or even uncover hidden physical objects in the scene (please see the sitting motion in the video starting from 3:09).

(R1) Q4: Multiplicity of solutions. A: For RFC-Explicit it is indeed possible to have multiple solutions since there are excessive DoFs when using multiple residual forces. However, RFC-Implicit introduces minimal new action dimensions and the regularization reward further reduces the set of feasible solutions. The question on multiplicity of solutions also applies to imitation learning in general since its only difference with RFC-Implicit is the root actuation.

(R2) Q5: Residual forces not in the original agent design. A: For human motion imitation, we are exactly trying to let an agent do something it is not able to do using the residual forces (RFs), since the agent is much simpler than real humans. If we don't want the agent to go beyond its ability, then RFC could be extended to a scaffolding technique (please see Q3). Also, as shown in the video, when the agent is forced to imitate demonstrations from other agents (e.g., human), the learned motions without RFs won't be more real. Instead, the agent often fails to imitate the motion or just falls to the ground. Finally, for agent-object interaction, the RFs won't hinder learning since the policy can always learn to output 0 RFs if they don't help. The RFs are only applied to stabilize the agent without changing object contact.

(R2) Q6: Additional evaluation. A: Since the motion synthesis baselines are deterministic, i.e., no diversity (we choose them because they allow stable long-term prediction), it would be unfair to compare diversity with them. The diversity of RFC is similar to our cVAE kinematic model (e.g., 5.7 (cVAE) vs. 5.6 (RFC) for average pairwise sample distance). Besides, the design of the cVAE itself is not the focus of the paper and can be replaced by other models.

(R3) Q7: Relation to CIO. A: These works are indeed related and we will include a discussion. However, we'd like to stress three key differences: (1) The residual forces (RFs) are **not** the contact forces (CFs) in CIO. RFs are **residuals** to existing CFs in the simulation which makes them easier to learn and directly quantify the physics violation. (2) Learning a policy that outputs RFs with RL is novel which can generalize to test data as shown in the motion synthesis experiments, i.e., no need to solve a trajectory optimization again. (3) Our RFC-Implicit formulation differs from the CIO physics regularization in that it isolates the physics violation to the root actuation (physically-valid w/o it). This again shows RFs differ from CFs since RFs are exact changes to physics decoupled from existing CFs in the simulation.

(R3) Q8: Comparisons with directly controlling state deltas. A: We did try this approach but it has a major problem: the delta changes in states do not go through physics simulation and often break contact, leading to unstable simulation.

(R3) Q9: Robustness of the learned policy. A: During training the RL policy injects Gaussian noise into the states, which allows the policy to explore nearby states around the demonstration as well, so the policy is robust to noise.

(R4) Q10: Comparison with MCP. A: MCP is a hierarchical method to combine different primitive skills to achieve tasks. As mentioned in the MCP paper, the primitive skills are still learned through DeepMimic. We used DeepMimic as baseline since it provides a pure comparison for motion imitation without the extra level of complexity.

(R4) Q11: No hard constraints on learned forces. A: The learned residual forces (RFs) are not to replace the contact forces (please see Eq. 1) but to correct the contact forces and compensate for the dynamics mismatch. So it is still valid for the RFs to be negative since the sum of the RFs and contact forces won't be negative. Further, the CoM forces are injecting energy only when they are needed to compensate for the dynamics mismatch (e.g., incorrect mass distribution) and match the demonstration. Finally, we can also impose hard constraints on the RFs by thresholding their magnitude.